

## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

This Page Blank (uspto)

OM protein - protein search, using sw model							
Run on:	April 18, 2005, 19:35:54 ; Search time 13.5116 Seconds (without alignments)						
Scoring table:	BLOSUM62 Gapop 10.0 , Gapext 0.5						
Searched:	283416 seqs, 96216763 residues						
Total number of hits satisfying chosen parameters:	283416						
Minimum DB seq length:	0						
Maximum DB seq length:	2000000000						
Post-processing:	Maximum Match 0% Listing First 45 summaries						
Database :	PIR 79;* 1: p1r1;* 2: p1r2;* 3: p1r3;* 4: p1r4;*						
Description	----- hypothetical prote hypothetical prote probable ABC trans hypothetical prote phosphatidate cyti hypothetical prote opha protein limpo hypothetical prote conserved hypothet cytochrome c bioge probable hexosyltr probable hexosyltr probable ABC trans cytochrome c-type probable cytochrom probable sodium-de hypothetical prote probable peptide/a hypothetical prote histidine transpor amino acid transpo hypothetical prote sucrose alpha-gluc IG heavy chain pre ammonia monooxygen						
Result No.	Score	Query	Match	Length	DB	ID	Description
1	54	100.0	83	2	E69903		hypothetical prote
2	54	100.0	187	2	GB1047		hypothetical prote
3	49	90.7	273	2	E92268		probable ABC trans
4	45	83.3	68	2	A11956		hypothetical prote
5	45	83.3	271	2	FB188		phosphatidate cyti
6	45	83.3	271	2	JC4832		hypothetical prote
7	42	77.8	563	2	JB2975		opha protein limpo
8	42	77.8	563	2	C98307		hypothetical prote
9	42	77.8	760	2	T41644		trp-a
10	41	75.9	280	2	P81984		hypothetical prote
11	41	75.9	280	2	HB1038		conserved hypothet
12	41	75.9	312	2	C70475		cytochrome c bioge
13	41	75.9	357	2	G66290		probable hexosyltr
14	41	75.9	400	1	F69442		probable hexosyltr
15	41	75.9	536	2	G93389		probable ABC trans
16	41	75.9	664	2	B75532		cytochrome c-type
17	41	75.9	699	2	F72453		probable cytochrom
18	40	74.1	211	2	S08522		probable sodium-de
19	40	74.1	448	2	T23263		hypothetical prote
20	40	74.1	545	2	A84432		probable peptide/a
21	40	74.1	568	2	E96648		hypothetical prote
22	40	74.1	595	2	C94432		histidine transpor
23	40	74.1	586	2	S46236		amino acid transpo
24	40	74.1	602	1	S38111		hypothetical prote
25	40	74.1	1039	2	T28905		sucrose alpha-gluc
26	40	74.1	1827	1	UUHU		IG heavy chain pre
27	40	74.1	1841	2	T10799		ammonia monooxygen
28	39	72.2	142	2	C34903		
29	39	72.2	276	2	A82550		



102 GYWGGRW 108

A; Reference number: Z222006

C;Superfamily: hypothetical protein HI037

Query Match    3 WGGYW 7  
 Best Local Similarity    75.9%; Score 41; DB 2; Length 280;  
 Matches    5; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy    3 WGGYW 7  
 Db    99 WGGYW 103

RESULT 12

C70475

Cytchrome c biogenesis protein - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C;Accession: C70475

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Residues: 1-312 &lt;AQF&gt;

A;Cross-references: UNIPROT:O67831; GB:AE000769; NID:92984262; PIDN:AAUC07795.1; PMID:92984262

A;Experimental source: strain VPS

A;Gene: hemX2

C;Superfamily: cytochrome c-type synthesis protein

Query Match    3 WGGYW 7

Best Local Similarity    75.9%; Score 41; DB 2; Length 312;  
 Matches    5; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy    3 WGGYW 7  
 Db    244 WGGYW 248

RESULT 13

G69290

probable hexosyltransferase (EC 2.4.1.-) AF0327 [similarity] - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: G69290

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 384-370, 1997

A;Authors: Utterback, T.; Corton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: G69290

A;Status: nucleic acid sequence not shown; translation not shown

A;Residues: 1-357 &lt;KE&gt;

A;Cross-references: UNIPROT:O29920; GB:AE001082; NID:92689405; PIDN:AAB9090

C;Superfamily: probable hexosyltransferase YtxN

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match    2 YWGGYW 7

Best Local Similarity    75.9%; Score 41; DB 2; Length 357;  
 Matches    5; Conservative    0; Mismatches    1; Indels    0; Gaps    0;

Qy    2 YWGGYW 7  
 Db    120 YWGDYW 125

F69142 probable hexosyltransferase (EC 2.4.1.-) MTH332 [similarity] - Methanobacterium thermoautotrophicum related protein

N;Alternative names: LPS biosynthesis rfb related protein

C;Species: Methanobacterium thermoautotrophicum

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: F69142

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanji, N.; Kir, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neilling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: F69142

A;Status: nucleic acid sequence not shown; translation not shown

C;Keywords: glycosyltransferase YtxN

A;Cross-references: UNIPROT:O26432; GB:AE000818; NID:92621384; PIDN:AAB8428

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH332

C;Superfamily: probable hexosyltransferase YtxN

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match    2 YMGGYW 7

Best Local Similarity    75.9%; Score 41; DB 1; Length 400;  
 Matches    5; Conservative    0; Mismatches    1; Indels    0; Gaps    0;

Qy    2 YMGGYW 7  
 Db    155 YMWDYW 160

RESULT 15

G55389 probable ABC transporter, periplasmic solute-binding protein SMA1860 [imported]

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C;Accession: G55389

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barlow-Hubler, F.; Bowe, R.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Weil, S.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A;Reference number: A95262; MUID:21396509; PMID:11481432

A;Status: Preliminary

A;Accession: G55389

A;Molecule type: DNA

A;Cross-references: UNIPROT:O92Y64; GB:AE006469; PIDN:AAK65681.1; PMID:914524171; GSPDB:&lt;

A;Experimental source: strain 1021, megaplasmid pSvMA

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Delai, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreanno, S.; Federici, N.A.; Fisher, R.F.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMA1860

A;Genome: plasmid

Query Match    1 GYWGYY 7

Best Local Similarity    75.9%; Score 41; DB 2; Length 536;  
 Matches    5; Conservative    0; Mismatches    2; Indels    0; Gaps    0;

Qy    1 GYWGYY 7  
 Db    417 GWDNYW 423

Search completed: April 18, 2005, 19:51:53  
Job time : 15.5116 secs

RESULT 14



*This Page Blank (uspto)*

Scoring table:	BLOSUM62	Alignments		
Gapopen:	10.0	Gapext 0.5		
Searched:	1612378 seqs, 512079187 residues			
Total number of hits satisfying chosen parameters:	1612378			
Minimum DB seq length:	0			
Maximum DB seq length:	2000000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
Database :	UniProt_03: 1: uniprot_sprot: 2: uniprot_trembl: *			
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
	SUMMARIES			
Result No.	Query Score	Match Length	DB ID	Description
1	54	100.0	83	1 YODI_BACSU
2	54	100.0	187	2 Q9HVT5
3	54	100.0	189	2 Q8QQB8
4	54	100.0	440	2 Q6VU6
5	49	90.7	273	2 Q93OZ1
6	49	90.7	273	2 Q89KL1
7	48	89.9	296	2 Q82LB6
8	47	87.0	363	2 Q8U043
9	46	85.2	214	2 Q8ABU4
10	45	83.3	68	2 Q8YXK8
11	45	83.3	106	2 Q9VZB1
12	45	83.3	271	1 CDSA_SEAE
13	45	83.3	762	2 Q7SD72
14	44	81.5	302	2 Q6LSS2
15	42	77.8	210	2 Q88RD4
16	42	77.8	261	2 CAPA_ASHGO
17	42	77.8	273	2 Q81A84
18	42	77.8	326	2 Q81A83
19	42	77.8	368	2 Q73TG1
20	42	77.8	441	2 Q7NB86
21	42	77.8	563	2 Q8UAG3
22	42	77.8	760	2 Q94533
23	42	77.8	982	2 Q64YTB
24	42	77.8	982	2 Q8A1Z4
25	41.5	76.9	429	2 Q8AAZ7
26	41.5	76.9	462	1 NORM_BURVI
27	41.5	76.9	468	2 Q62LW6
28	41.5	76.9	468	2 Q63SS57
29	41	75.9	236	2 Q63TP2
30	41	75.9	280	2 Q50469
31	41	75.9	280	2 Q9Jvw6
				neisseria m
				Q9jy00 neisseria m
				Q62K29 burkholderi
				Q67831 aquifex aeo
				Q29920 archaeoglob
				Q7NpV4 chromobacte
				Q8xv46 ralstonia s
				Q9ad5 streptomyce
				Q62gn8 burkholderi
				Q63q5 burkholderi
				Q26432 methanobact
				Q8tL58 methanosarc
				Q98t18 rhizobium 1
				Q6w118 rhizobium 8
				Q92y64 rhizobium m

RA Vassarotti A., Viari A., Wambutt R., Wedler B., Wedler H.,  
 RA Weitznecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zunstein E.,  
 RA Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis";  
 RL Nature 390:249-256 (1997).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to licensee@isb-sib.ch).  
 CC  
 DR EMBL; AF015775; AAB72056.1; -.  
 DR EMBL; AF006665; AAB81166.1; -.  
 DR EMBL; Z99114; CAB13852.1; -.  
 DR PIR; B69903; B69903; B69903.  
 DR SubLibList; BG13537; YodI.  
 DR InterPro; IPR008991; Transl SH3-like.  
 DR Complete proteome; Hypothetical protein; Transmembrane.  
 KW TRANSMEM 58 80 Potential.  
 FT SEQUENCE 83 AA; 9194 MW; 99F8EA22E0F36A43 CRC64;  
 SQ Query Match 100.0%; Score 54; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GYWGGYW 7  
 Db 53 GYWGGYW 59

RESULT 2  
 Q9HV15 PRELIMINARY; PRT; 187 AA.  
 AC Q9HV15  
 ID Q9HV15  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=PA4793;  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NCBI\_TaxID=287; RN [1]  
 SEQUENCE FROM N\_A.  
 RC STRAIN=ATCC 15692 / PA01; ID=10\_103B/3502/079;  
 RX MEDLINE-40437337; PubMed=10384043; DOI=10.103B/3502/079;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.J., Mizoguchi S.D., Warrener P.,  
 RA Hickey M.J., Brinkman F.S.L., Hurnagie W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Collier L.R., Tolentino E., Westbroek-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Polger K.B., Kas A., Verbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.W.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an  
 opportunistic pathogen.";  
 RT Complete proteome; Hypothetical protein.  
 DR EMBL; AP004892; AACG08179.1; -.  
 DR PIR; GB3047; GB3047.  
 SQ SEQUENCE 187 AA; 21281 MW; 8908E8EBEBC51897 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWGGYW 7  
 Db 109 GYWGGYW 115

QY	1 GYWGYYW 7				RX MEDLINE=22484998; PubMed=12597275;
Db	212 GYWGYYW 218				RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
					RA Sasamoto S., Watanabe A., Idesawa K., Iriuguchi M., Kawashima K.,
					RA Kohara M., Matsumoto M., Shimpo S., Tsuriuoka H., Wada T., Yamada M.,
					RA Tabata S.;
					RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110.";
					RL DNA Res. 9:189-197(2002)
					-I - FUNCTION: Part of a binding-protein-dependent transport system across the membrane (By similarity).
					CC -I - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
					CC -I - SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.
					CC EMBL; AP00553; BAC5015.1; -.
					DR GO; GO:0016021; C:integral to membrane; IEA.
					DR GO; GO:0005215; F:transporter activity; IEA.
					DR GO; GO:0006810; P:transport; IEA.
					DR InterPro; IPR000515; BPD transp.
					DR Pfam; PF00528; BPD transp_1; 1.
					DR PROSITE; PS50928; ABC_TMI; 1.
					KW Complete proteome; Transmembrane; Transport;
					SQ SEQUENCE 273 AA; 29042 MW; 6E3D8EE197497613 CRC64;
					Query Match 90.7%; Score 49; DB 2; Length 273;
					Best Local Similarity 85.7%; Pred. No. 17; Mismatches 0; Indels 0; Gaps 0;
					Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
					Qy 1 GYWGYYW 7
					Db 94 GYWGWW 100
					RESULT 7
					Q8ZLB6 PRELIMINARY; PRT; 296 AA.
					ID Q8ZLB6
					AC Q8ZLB6;
					DT 01-JUN-2003 (TrEMBLrel. 24, Created)
					DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
					DR 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
					DR Purative sugar transporter.
					GN Orderidocusnames=SAV2094;
					OS Streptomyces avermitilis.
					OC Bacteria; Actinobacteria; Actinomycetales;
					OC Streptomyicinae; Streptomyctaceae; Streptomyces.
					OX NCBI_TaxID=33903;
					RN [1]_TAXID=33903;
					RP SEQUENCE FROM N.A.
					RC STRAIN=MA-4580;
					RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
					RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
					RA Shiose M., Takahashi Y., Horikawa H., Nakazawa H., Osonee T.,
					RA Kikuchi H., Shiba T., Sakai Y., Hattori M.,
					RT "Genome sequence of an industrial microorganism Streptomyces avermitilis"; deducing the ability of producing secondary metabolites";
					Proc. Natl. Acad. Sci. U.S.A. 98:12225-12220(2001).
					RL RN [2]
					RP SEQUENCE FROM N.A.
					RC STRAIN=MA-4680;
					RC MEDLINE=22603306; PubMed=12692562;
					RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
					RA Sakai Y., Hattori M., Omura S.;
					RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.";
					RI RT Natl. Biotechnol. 21:526-531(2003).
					CC -I - FUNCTION: Part of a binding-protein-dependent transport system across the membrane (By similarity).
					CC -I - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
					CC -I - SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.
					CC EMBL; AP005029; BAC69805.1; -.

DR	GO; GO:0016021; C:integral to membrane; IEA.	RT	"A genomic view of the human-Bacteroides thetaiotaomicron symbiosis." RL Science 299(2074-2076)(2003).
DR	GO; GO:005351; F:sugar porter activity; IEA.	RL	DR EMBL; AE01630 ; AA076180.1 ; - .
DR	GO; GO:0005215; F:transporter activity; IEA.	KW	Complete proteome.
DR	Interpro; IPR006810; P:transport; IEA.	SEQUENCE	214 AA; 24735 MW; 84ABB75A740D226B CRC64;
Pfam	Pfam0528; BPD transp_1; 1.	Query	Match 85.2%; Score 46; DB 2; Length 214;
DR	PROSITE; PS50928; ABC_TML_1.	Best Local Similarity 85.7%; Pred. No. 38;	
DR	Complete proteome; sugar transport; Transmembrane; Transport.	Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
SEQUENCE	296 AA; 32715 MW; 70374AD2BB8F8FB8A CRC64;	Qy	1 GYNGGYW 7
Qy	1 GYNGGYW 7	Db	130 GRGPGYW 136
Db	198 GYGAYW 204	RESULT 10	
		08YXXKB	PRELIMINARY; PRT; 68 AA.
		ID Q8YXXKB	
		AC Q8YXXKB	
		DT 01-MAR-2002 (TREMBLrel. 20; Created)	
		DT 01-MAR-2002 (TREMBLrel. 20; Last sequence update)	
		DT 01-JUN-2003 (TREMBLrel. 24; Last annotation update)	
		DE Apr204 protein.	
		DN OrderedLocusNames=asr1204;	
		OS Anabaena sp. (strain PC 7120).	
		RA Bacteria; Cyanobacteria; Nostocales; Nostoc.	
		[1] NCBI_TaxID=103690;	
		RN SEQUENCE FROM N.A.	
		RP MEDLINE=21595285; PubMed=11759840;	
		RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,	
		RA Watanabe A., Iriuchi M., Ishikawa K., Kimura T.,	
		RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,	
		RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,	
		RA Yasuda M., Tabata S.;	
		RT "Complete genomic sequence of the filamentous nitrogen-fixing	
		RT cyanobacterium Anabaena sp. strain FCC 7120."	
		RT DNA Res. 8:205-213 (2001).	
		DR EMBL: AP003585; BAB73161.1; - .	
		DR PR; AI1956; AI1956.	
		KW Complete proteome.	
		SEQUENCE 68 AA; 7944 MW; BE31A1488358C962 CRC64;	
		Query Match 83.3%; Score 45; DB 2; Length 68;	
		Best Local Similarity 85.7%; Pred. No. 20;	
		Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
		Qy	1 GYNGGYW 7
		Db	61 GRGPGYW 67
RESULT 11		Q9VZB1	PRELIMINARY; PRT; 106 AA.
		ID Q9VZB1	
		AC Q9VZB1	
		DT 01-MAY-2000 (TREMBLrel. 13; Created)	
		DT 01-OCT-2002 (TREMBLrel. 13; Last sequence update)	
		DE CG13721-PA.	
		GN ORFNames=CG13721;	
		OS Drosophila melanogaster (Fruit fly).	
		OC Drosophila; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
		OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
		OC Ephydriidae; Dirosophilidae; Drosophila.	
		OX NCBI_TaxID=7227;	
		RN [1] SEQUENCE FROM N.A.	
		RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;	
		RC Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
		RC Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
		Xu J., Bjurzell M.K., Himrod J., Deng S., Richards S., Ashburner M., Carmichael L.K.,	
		RA Chiang H.C., Hooper L.V., Gordon J.I., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	

RA	Brandon R.C., Rogers Y.H., Blazej R.G., Champé M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Hell G., Nelson C.R., Gabor G.L., Abril J.F., Agayani A., An H.J., Andrews-Pflanck C., Baldwin G., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Beno P.V., Berman B.P., Bhandari D., Bolhakov S., Borkova D., Botchan M.V., Bouck J., Brostean P., Brottier P., Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dorson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferrara S., Fleischmann W., Foster C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K., Glodzik A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nixon K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puriv V., Reese M.G., Reiner K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svartkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Ye J., Yen R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., "The genome sequence of <i>Drosophila melanogaster</i> ," <i>Science</i> 287:2185-2195 (2000).	RP	SEQUENCE FROM N.A.
[2]	RP	SEQUENCE FROM N.A.	
RA	Patel S., Adams M., Changé M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; "Finishing a whole-genome shotgun: Release 3 of the <i>Drosophila melanogaster</i> euchromatic genome sequence.," <i>Genome Biol.</i> 3:RESEARCH0079-RESEARCH0079 (2002).	RP	SEQUENCE FROM N.A.
RA	Patel S., Adams M., Changé M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; "Finishing a whole-genome shotgun: Release 3 of the <i>Drosophila melanogaster</i> euchromatic genome sequence.," <i>Genome Biol.</i> 3:RESEARCH0079-RESEARCH0079 (2002).	RP	SEQUENCE FROM N.A.
[3]	RP	SEQUENCE FROM N.A.	
RA	Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celinker S.E.; "The transposable elements of the <i>Drosophila melanogaster</i> euchromatin: a genomic perspective.," <i>Genome Biol.</i> 3:RESEARCH0084-RESEARCH0084 (2002).	RP	SEQUENCE FROM N.A.
RA	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., RA Bettencourt B.R., Celinker S.E., Whittied B.J., Bayraktaroglu L., Berman B.P., RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., RA Lewis S.E.; "Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a RA genomic review.," <i>Genome Biol.</i> 3:RESEARCH0083-RESEARCH0083 (2002).	RP	SEQUENCE FROM N.A.
[4]	RP	SEQUENCE FROM N.A.	
RA	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., RA Bettencourt B.R., Celinker S.E., Whittied B.J., Bayraktaroglu L., Berman B.P., RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., RA Lewis S.E.; "Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a RA genomic review.," <i>Genome Biol.</i> 3:RESEARCH0083-RESEARCH0083 (2002).	RP	SEQUENCE FROM N.A.
[5]	RP	SEQUENCE FROM N.A.	
RA	FlyBase; Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.	RG	FlyBase;
RA	Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.	RL	Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN		RN	
RA	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RA	FlBase; Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.	RG	FlBase;
RL	EMBL; AE003481; AAFA7913.1; -	DR	EMBL; AE003481; AAFA7913.1; -
DR	Inexact; Q9VZB1; -	DR	Inexact; Q9VZB1; -
DR	FlBase; FB9003554; CG13721.	DR	FlBase; FB9003554; CG13721.
SQ	SEQUENCE 106 AA; BA74FF7C320474EE CRC64;	SQ	SEQUENCE 106 AA; BA74FF7C320474EE CRC64;
Query Match	83.3%; Score 45; DB 2; Length 106;	Query Match	83.3%; Score 45; DB 2; Length 106;
Best Local Similarity 85.7%; Pred. No. 29;	Best Local Similarity 85.7%; Pred. No. 29;	Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 GYWGGSYW 7	QY	1 GYWGGSYW 7
Db	68 GYWGGSYW 74	Db	68 GYWGGSYW 74
RESULT 12	CDS,_PSRAE STANDARD; PRT; 271 AA.	RESULT 12	CDS,_PSRAE STANDARD; PRT; 271 AA.
ID	Q56440; AC	ID	Q56440; AC
DT	01-NOV-1997 (Rel. 35, Created)	DT	01-NOV-1997 (Rel. 35, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)	DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	05-JUL-2004 (Rel. 44, Last annotation update)	DT	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diacylglyceride synthetase) (CDP-diacylglyceride pyrophosphorylase) (CDP-diacylglycerol synthase) (CDS) (CDP:phosphate diacyldiyltransferase) (CDP-DAG synthase) (CDP-DG synthetase).	DE	Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diacylglyceride synthetase) (CDP-diacylglyceride pyrophosphorylase) (CDP-diacylglycerol synthase) (CDS) (CDP:phosphate diacyldiyltransferase) (CDP-DAG synthase) (CDP-DG synthetase).
GN	Name=cddb; Synonyms=cdd; OrderedlocusName=PA3651;	GN	Name=cddb; Synonyms=cdd; OrderedlocusName=PA3651;
OS	Pseudomonas aeruginosa	OS	Pseudomonas aeruginosa
OC	Bacteria: Proteobacteria: Gammaproteobacteria: Pseudomonadales: Pseudomonadaceae: Pseudomonas.	OC	Bacteria: Proteobacteria: Gammaproteobacteria: Pseudomonadales: Pseudomonadaceae: Pseudomonas.
OX	NCBITaxonID=287;	OX	NCBITaxonID=287;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 15692 / PACO1; MEDLINE=22426065; PubMed=12537568;	RC	STRAIN=ATCC 15692 / PACO1; MEDLINE=22426065; PubMed=12537568;
RX	MEDLINE=62257274; PubMed=8654980; DOI=10.1016/0378-1119(96)00009-1;	RX	MEDLINE=62257274; PubMed=8654980; DOI=10.1016/0378-1119(96)00009-1;
RA	Taguchi K., Fukutomi H., Kuroda A., Kato J., Ontake H.; Cloning of the <i>Pseudomonas aeruginosa</i> gene encoding CDP-diacylglyceride synthetase.," <i>J. Bacteriol.</i> 172:165-166 (1996) .	RA	Taguchi K., Fukutomi H., Kuroda A., Kato J., Ontake H.; Cloning of the <i>Pseudomonas aeruginosa</i> gene encoding CDP-diacylglyceride synthetase.," <i>J. Bacteriol.</i> 172:165-166 (1996) .
RT	"	RT	"
RN	[2]	RN	[2]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 15692 / PACO1; MEDLINE=0431337; PubMed=10384043; DOI=10.1038/35023079;	RC	STRAIN=ATCC 15692 / PACO1; MEDLINE=0431337; PubMed=10384043; DOI=10.1038/35023079;
RX	Hickey M.J., Brinkman P.S.L., Huffnagle B.W., Kowalski D.J., Lagrou M., Gaber R.L., Goltry L., Tolentino E., Westrock-Wadman S., Yuan Y., Brody L.L., Coulter S.H., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Pulsen I.T., Reizer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.;	RX	Hickey M.J., Brinkman P.S.L., Huffnagle B.W., Kowalski D.J., Lagrou M., Gaber R.L., Goltry L., Tolentino E., Westrock-Wadman S., Yuan Y., Brody L.L., Coulter S.H., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Pulsen I.T., Reizer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA	"Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen." <i>Nature</i> 406:959-964 (2000); RT RT	RA	"Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen." <i>Nature</i> 406:959-964 (2000); RT RT
RL	Nature 406:959-964 (2000); CC -1- CATALYTIC ACTIVITY: CTP + phosphatidate = diophosphate + CDP-diacylglycerol.	RL	Nature 406:959-964 (2000); CC -1- CATALYTIC ACTIVITY: CTP + phosphatidate = diophosphate + CDP-diacylglycerol.
CC	-1- PATHWAY: Phospholipid biosynthesis. CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.	CC	-1- PATHWAY: Phospholipid biosynthesis. CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC	-1- SIMILARITY: Belongs to the CDS family.	CC	-1- SIMILARITY: Belongs to the CDS family.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
DR	EMBL; D50811; BAA09437.1; -	DR	EMBL; D50811; BAA09437.1; -
DR	EMBL; AE004785; AAG07039.1; -	DR	EMBL; AE004785; AAG07039.1; -
DR	PIR; F83188;	DR	PIR; F83188;
DR	PIR; JC4832;	DR	PIR; JC4832;
DR	InterPro; IPR000374; PC_trans.	DR	InterPro; IPR000374; PC_trans.

DR pfam; PF01148; CTP\_transf\_1; 1.  
 DR PROSITE; PS01315; CDPs; 1.  
 KW Complete proteome; Inner membrane; Nucleotidyltransferase;  
 KW Phospholipid biosynthesis; Transferase; Transmembrane.  
 FT TRANSEM 12 32 Potential.  
 FT TRANSEM 53 73 Potential.  
 FT TRANSEM 75 95 Potential.  
 FT TRANSEM 111 131 Potential.  
 FT TRANSEM 136 156 Potential.  
 FT TRANSEM 174 194 Potential.  
 FT TRANSEM 199 219 Potential.  
 FT TRANSEM 251 271 Potential.  
 FT CONFLICT 131 133 WPI->VAA (in Ref. 1).  
 SQ SEQUENCE 271 AA; 28856 MW; 5025059C3F1A64C7 CRC64;  
 Query Match 83.3%; Score 45; DB 1; Length 271;  
 Best Local Similarity 85.7%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYWGTYW 7  
 Db 102 GYWGGRW 108

RESULT 13  
 ID Q7SD72 PRELIMINARY; PRT; 762 AA.  
 AC DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 AC DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 AC DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Predicted protein.  
 GN Name=NCU02790.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 RN [1] \_TaxID=5141;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Or74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D.T., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins C., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburn M.,  
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zaitsev A., Schulze U.,  
 RA Rothe G.O., Jedd G., Maves W., Staben C., Marcott E., Greenberg D.,  
 RA Kamal M., Kamry Seelis M., Mauceli E., Bielek C., Rudd S., Frishman D.,  
 RA Krytofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,  
 RA Cogeni C., Macino G., Catchpole D., Li W., Pratt R.J., Osman S.A.,  
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yilmaz M., Seilzer S., Dunlap J., Berford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Manhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0-0(2003).  
 CC -!- ATTENTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAPX01000062; EAA34710.1; -.  
 DR InterPro; IPR011014; MscS\_transmembr.  
 SQ SEQUENCE 762 AA; 84704 MW; 2C7DD1515E1902FA CRC64;  
 Query Match 83.3%; Score 45; DB 2; Length 762;  
 Best Local Similarity 71.4%; Pred. No. 1.e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GYWGTYW 7  
 Db 554 GPWGQWW 560

RESULT 14  
 ID Q6LSS8 PRELIMINARY; PRT; 302 AA.  
 AC Q6LSS8; 27, Created)  
 AC DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 AC DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=PBPPA1237;  
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Photobacterium.  
 NCBI\_TaxID=74109;  
 RN [1] \_TaxID=74109;  
 RP SEQUENCE FROM N.A.  
 RA Verzi A., Campanaro S., D'Angelo M., Simonato F., Vitullo N., Lauro F.,  
 RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,  
 RA Vale G.;  
 RT "Genome analysis of Photobacterium profundum reveals the complexity of  
 RT high pressure adaptations.";  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; CP378667; CAG19648.1; -.  
 KW Complete genome.  
 SQ SEQUENCE 302 AA; 34200 MW; 708FAB0ED34BADE7 CRC64;

RESULT 15  
 ID Q88RD4 PRELIMINARY; PRT; 210 AA.  
 AC Q88RD4; 24, Created)  
 AC DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 AC DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 Qy 2 YMGGYW 7  
 ||||:|  
 Db 199 YMGGFW 204

RESULT 15  
 ID Q88RD4 PRELIMINARY; PRT; 210 AA.  
 AC Q88RD4; 24, Created)  
 AC DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 AC DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Transporter, Lyse family.  
 GN OrderedLocusNames=PP0198;  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NCBI\_TaxID=160488;  
 RN [1] \_TaxID=160488;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouls D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L.M., Beanan M.J., Desby R.T., Daugerty S.C., Kolonay J.F.,  
 RA Madupu R., Nelson W.C., White P., Peterson J.D., Khouri H.M.,  
 RA Hance I., Chris Lee P., Holtzapffe E.K., Scanlan D., Tran K.,  
 RA Moazzze A., Utterback T.R., Rizzo M., Lee K., Koback D., Moestl D.,  
 RA Wedler H., Luber J., Stiepanik D.N., Hoheisel J., Streit J., Heim S.,  
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesteboef A., Tuemmler B.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808 (2002).  
 DR EMBL; AB016774; AAN65831.1; -.  
 DR TIGR; PP0198; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005293; Lysine permease activity; IEA.  
 DR GO; GO:0008865; F:amino acid transport; IEA.  
 DR InterPro; IPR001123; LysE.  
 DR Pfam; PF01810; LysE; 1.  
 KW Complete proteome.

Query Match 77.8%; Score 42; DB 2; Length 210;  
 Best Local Similarity 71.4%; Pred. No. 1.e-02;

	Matches	5;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
QY		1	GYWGGYW	7						
			:							
Db		35	GEWRGYW	41						

Search completed: April 18, 2005, 19:58:23  
Job time : 65.3488 secs



Page  
1

ABO04396 represent *P. aeruginosa* polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdb.uspto.gov/sequence.html

XX	Sequence 339 AA;
QQ	Query Match ID ABU20398
CC	Best Local Similarity 100.0%;
CC	Matches 7; Conservative 0;
CC	Mismatches 0;
CC	Indels 0;
CC	Gaps 0
Db	1 GYWGGW 7
Db	261 GYWGGYW 267

**RESULT 2**

ABU20398	Query Match ID ABU20398 standard; protein; 440 AA.
XX	AC ABU20398;
XX	DT 19-JUN-2003 (first entry)
DE	Protein encoded by Prokaryotic essential gene #5925.
XX	KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	OS <i>Bacteroides fragilis</i> .
XX	PN WO200277183-A2.
XX	PD 03-OCT-2002.
XX	PP 21-MAR-2002; 2002WO-US009107.
XX	PR 21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-FEB-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-036699P.
XX	PA (BLT-) ELITRA PHARM INC.
XX	PT Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI	Wall D, Trwick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
PI	DR WPI; 2003-039926/02.
XX	DR N-PSDB; ACA24268.
PS	SEQ ID NO 48322; 1766pp; English.
XX	PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
XX	Claim 25; SEQ ID NO 48322; 1766pp; English.
XX	The invention relates to an isolated nucleic acid comprising any one of the 623 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibit proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX	SQ	Sequence 106 AA;	Db	331 GYWGRW 337
Query Match 83.3%; Score 45; DB 4; Length 106;	RESULT 5			
Best Local Similarity 85.7%; Pred. No. 36;	AY03714 standard; peptide; 8 AA.			
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	XX			
QY 1 GYNGGYW 7	XX			
Db 68 GYYGCGYW 74	XX			
RESULT 4	Fluorine-18 (F-18) labeled peptide 1.			
ABO72468 ID ABO72468 standard; protein; 500 AA.	XX			
AC ABO72468;	XX			
DT 29-JUL-2004 (first entry)	XX			
DE Pseudomonas aeruginosa polypeptide #4643.	XX			
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.	XX			
OS Pseudomonas aeruginosa.	XX			
PN US6551785-B1.	XX			
PD 22-APR-2003.	XX			
PF 18-FEB-1999; 99US-002522991.	XX			
PR 18-FEB-1998; 98US-007478BP.	XX			
PR 27-JUL-1998; 98US-0034190P.	XX			
PA (GENO-) GENOME THERAPEUTICS CORP.	XX			
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;	XX			
PS WPI; 2003-615309/58.	XX			
DR N-PSDB; ABD06039.	XX			
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of bacterial infection.	XX			
PT Disclosure: SEQ ID NO 21214; 455PP; English.	XX			
PS WPI; 1999-228967/19.	XX			
PT Radiolabeling thiol-containing peptides with fluorine-18.	XX			
CC The invention relates to Pseudomonas aeruginosa polypeptides and the CC poly nucleotides encoding them. The sequences are useful in diagnosis and CC therapy of pathological conditions, as molecular targets for diagnostics, CC poly nucleotides encoding them. The sequences are useful in diagnosis and CC therapy of pathological conditions, as molecular targets for diagnostics, CC for the ability to bind a P. aeruginosa nucleic acid, as components of CC effective antibiotic targets, as targets for antibacterial drugs, CC including anti-P. aeruginosa drugs, as templates for recombinant CC production of P. aeruginosa-derived peptides or polypeptides, as target CC components for diagnosis and/or treatment of P. aeruginosa-caused CC infection, and in detection of P. aeruginosa sequences or other sequences CC of Pseudomonas species using biochip technology. Sequences AB067822- CC AB067826 represent P. aeruginosa polypeptides of the invention. Note: The CC sequence data for this patent did not form part of the printed CC specification but was obtained in electronic format from USPTO at CC seqdata.uspto.gov/sequence.html	XX			
XX Sequence 500 AA;	XX			
SQ Query Match 83.3%; Score 45; DB 7; Length 500;	XX			
Best Local Similarity 83.7%; Pred. No. 1.5+02;	XX			
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	XX			
QY 1 GYNGGYW 7	XX			

Page 4

CC the target site and a capture arm that binds to a polymer conjugate, and  
 CC administering a polymer conjugate that binds to the capture arm, the  
 CC conjugate has a polymer conjugated to the agent such as therapeutic  
 CC agent, a peptide, an enzyme and a labelled ligand. Also included is a kit  
 CC useful for targeting a target site within a tissue in a subject or tissue  
 CC sample comprising the above mentioned multi-specific antibody or its  
 CC fragment and a polymer conjugate. The method is used for targeting an  
 CC agent towards a target site in a tissue (e.g. a tumour). The method is  
 CC also useful for therapeutic or diagnostic purposes and further in  
 CC photodynamic therapy. The present sequence is an Iodinated immunogenic  
 CC peptide used to raise antibodies which recognise only the iodinated  
 CC peptide, used in the method of the invention.

XX

SQ

Sequence 8 AA;

Query Match	80.6%	Score 43.5;	DB 7;	Length 8;
Best Local Similarity	87.5%	Pred. No. 1.8e+06;	Indels 0;	Gaps 1;
Matches	7;	Conservative	Mismatches 0;	

antibody having a targeting arm that binds to an antigen and a capture arm that binds to a polymer conjugate comprising a therapeutic agent.

PT Disclosure; Page 16; 24pp; English.

PT The invention relates to a method of diagnosing or treating a disease or disorder. The method involves administering to a tissue a multi-specific antibody (I) or antibody fragment, comprising a targeting arm that binds to an antigen on the target site, and a capture arm that binds to a polymer conjugate, and administering to the tissue a polymer conjugate conjugated to a diagnostic or therapeutic agent. Also included is a method for photodynamic diagnosis or treatment of a disease or disorder; or intravascular or endoscopic method for diagnosing or treating a disease or disorder chosen from cancer (oesophageal, gastric, colonic, rectal, pancreatic, lung, breast, ovarian, urinary bladder, endometrial, cervical, testicular, renal, adrenal and liver cancer, solid tumour, B-cell malignancy or T-cell malignancy); cardiovascular lesion; an inflammatory disease; neurodegenerative disease; metabolic renal an infectious disease. The B-cell malignancy is chosen from indolent lymphatic leukaemia, acute lymphatic leukaemias, and multiple myeloma. The solid tumour is chosen melanoma, carcinoma (preuterine renal carcinoma, lung carcinoma, intestinal carcinoma, and stomach carcinoma), glioma and sarcoma. The cardiovascular lesion is chosen from infarct, clot, embolus, atherosclerotic plaque and ischaemia. The neurodegenerative disease is Alzheimer's disease. The metabolic disease is amyloidosis, where the antibody binds amyloid. The disease or disorder is displaced or ectopic normal tissue chosen from endometrium, thymus, spleen and parathyroid. The method can be used for normal tissue ablation, where the tissue is chosen from bone marrow and spleen. The disease or disorder is an autoimmune disease such as myasthenia gravis, lupus nephritis, lupus erythematosus, and rheumatoid arthritis, Class III autoimmune diseases such as immune-mediated thrombocytopenias, such as acute idiopathic thrombocytopenic purpura and chronic idiopathic thrombocytopenic purpura, dermatomyositis, Sjogren's syndrome, multiple sclerosis, Sydenham's chorea, myasthenia gravis, systemic lupus erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes, bullous pemphigoid, diabetes mellitus, Henoch-Schonlein's arteritis, streptococcal nephritis, erythema nodosum, Takayasu's arteritis, spondylitis, Goodpasture's syndrome, thromboangiitis obliterans, primary biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma, chronic active hepatitis, polymyositis/dermatomyositis, polychondritis, pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy, pernicious anaemia, rapidly progressive arteritis/polymyalgia, pernicious anaemia, fibrosing alveolitis. The infectious disease is chosen from bacterial, fungal, parasitic and viral lesion. The infectious disease is caused by a fungus chosen from Microsporum, Trichophyton, Epidermophyton, Sporothrix schenckii, Cryptococcus neoformans, Coccidioides immitis, Histoplasma capsulatum, Blastomyces dermatitidis, and Candida albicans. The infectious disease is caused by a virus chosen from HIV, herpes virus, cyrmegalovirus, rabies virus, influenza virus, hepatitis B virus, sendai virus, feline leukemia virus, reo virus, polio virus, human serum parvo-like virus, simian virus 40, respiratory syncytial virus, mouse mammary tumour virus, Varicella-Zoster virus, Dengue virus, rubella virus, measles virus, adenovirus, human T-cell leukaemia viruses, Epstein-Barr virus, murine leukemia virus, mumps virus, vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis virus, wart virus and blue tongue virus. The infectious disease is caused by a bacterium chosen from Bacillus anthracis, Streptococcus agalactiae, Legionella pneumophila, Streptococcus pyogenes, Escherichia coli, Neisseria gonorrhoeae, Neisseria meningitidis, Pneumococcus, Haemophilus influenzae B, Treponema pallidum, Lyme disease spirochetes, Peudomonas aeruginosa, Mycobacterium leprae, Brucella abortus, Mycobacterium tuberculosis, and Tetanus toxin. The infectious disease is caused by a protozoa chosen from Plasmodium falciparum, Plasmodium vivax, Toxoplasma gondii, Trypanosoma rangeli, Trypanosoma cruzi, Trypanosoma rhodesiense, Trypanosoma brucei, Schistosoma mansoni, Schistosoma japonicum, Babesia bovis, Elmeria tenella, Onchocerca volvulus,

CC Leishmania tropica, Trichinella spiralis, Onchocerca volvulus, Theileria parva, Taenia hydatigena, Taenia ovis, Taenia saginata, Echinococcus granulosus, and Mesocystoides corti. The infectious disease is caused by a mycoplasma chosen from Mycoplasma arthritidis, M. hyorhinis, M. oralis, M. arginini, Acholeplasma laidlawii, M. salivarium and M. pneumoniae. The cancer is preferably chosen from carcinoma antigen (CEA) - expressing tumour or a CD20-expressing malignancy. The present sequence represents a peptide used in the method of the invention.

XX Sequence 8 AA;

Query Match	80.6%	Score 43.5;	DB 8;	Length 8;	
Best Local Similarity	87.5%;	Pred. No. 1.8e+06;			
Matches	7;	Conservative	0;	Mismatches	0;
Qy	1 GYWG-GYW 7				
Db	1 GYWGKGYW 8				

RESULT 9

AYT6816	DE Immunogenic peptide for bi-specific antibody recognition.
ID AAY76816	XX
AC AAY76816;	XX
XX DT 28-APR-2000 (first entry)	
XX DE Immunogenic peptide; bi-specific antibody; diagnosis; immune response;	
KW disease tissue identification; therapy.	
XX OS Synthetic.	
XX PH	
FT Key	
FT Misc-difference 1	/note= "free amino acid group, protected amino acid group, chelating agent or a metal-chelate complex"
FT Misc-difference 2	/note= "D-form residue"
FT Misc-difference 3	/note= "D-form residue"
FT Misc-difference 4	/note= "D-form residue"
FT Misc-difference 5	/note= "D-form residue; modified with free amino acid group, protected amino acid group, chelating agent or a metal-chelate complex"
FT Misc-difference 6	/note= "D-form residue"
FT Misc-difference 7	/note= "D-form residue"
FT Misc-difference 8	/note= "D-form residue"
FT Misc-difference 9	/note= "D-form residue"
XX PN WO9666951-A2.	
XX PD 29-DEC-1999.	
XX PF 22-JUN-1999;	99WO-US013879.
XX PR 22-JUN-1998;	98US-0090142P.
XX PR 14-OCT-1998;	98US-0104156P.
PA (IMMU-) IMMUNOMEDICS INC.	
XX PI Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;	
XX DR WPI; 2000-160561/14.	
XX PR Bi-specific antibodies that bind specific target tissue and targeted conjugates.	
PS Claim 21; Page 61; 76pp; English.	
XX This sequence represents an immunogenic peptide for a bi-specific	

CC antibody. The invention relates to a method of treating or identifying CC diseased tissues in a patient comprising administering a bi-specific CC antibody (or fragment) having at least 1 arm (A) that specifically binds CC a targeted tissue and at least 1 arm (B) that specifically binds a CC targetable conjugate. The methods and bi-specific antibodies and fusion CC proteins are useful for pre-targeting methods of diagnosis and therapy. CC It is advantageous to raise bi-specific antibodies against a targetable CC conjugate that is capable of carrying at least 1 diagnostic or CC therapeutic agent. The characteristics of the chelator, metal chelate CC complex, therapeutic agent or diagnostic agent can be varied to CC accommodate differing applications without raising new bi specific CC antibodies for each new application. The targetable conjugate is selected CC to elicit sufficient immune responses and also for rapid in vivo CC clearance when used within the bi-specific antibody targeting method CC

XX SQ Sequence 9 AA;

Query	Match	80.6%;	Score 43.5;	DB 3;	Length 9;
Best Local Similarity	87.5%;	Pred. No. 1.8e+06;	Mismatches	0;	Gaps 1;
Matches	7;	Conservative	0;	Indels	1;
Qy	1 GTWG-GTW 7				
Db	2 GTWGKGWT 9				

RESULT 10

AY011208	DE Human polypeptide SEQ ID NO 25100.
ID AY011208	XX standard; protein; 131 AA.
AC AY011208;	XX
XX DT 06-NOV-2001 (first entry)	
XX DE Human polypeptide SEQ ID NO 25100.	
XX ID AY011208	
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW nervous system disorders; arthritis; inflammation.	
XX OS Homo sapiens.	
XX PN WO200164835-A2.	
XX PR 07-SEP-2001.	
XX PR 26-FEB-2001; 2001WO-US004927.	
XX PR 28-FEB-2000; 2000US-00515126.	
PR 18-MAY-2000; 2000US-00577409.	
XX DR 2001-514838/56.	
XX PA (HYSE-) HYSEQ INC.	
XX PI Tang YT, Liu C, Drmanac RT;	
XX PR WPI; 2001-514838/56.	
XX DR N-P5DB; AY011208.	
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing CC and treating e.g. leukemia, inflammation and immune disorders.	
XX PR	
XX PS Claim 20; SEQ ID NO 25100; 1399pp + Sequence Listing; English.	
XX DR	
CC The invention relates to human polynucleotides (AAI79941-AAI93861) and CC the encoded Proteins (AA00010-AA013910) that exhibit activity relating to CC cytokine, cell proliferation or cell differentiation or which may induce CC production of other cytokines in other cell populations. The CC polynucleotides and polypeptides are useful in gene therapy, vaccines or CC peptide therapy. The polypeptides have various cytokine-like activities, CC e.g. stem cell growth factor activity, haemopoiesis regulating CC activity, tissue growth factor activity, immunomodulatory activity and CC activin/inhibin activity and may be useful in the diagnosis and/or	

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

SQ Sequence 131 AA;

Query Match	Score	DB	Length	131;
Best Local Similarity	79.6%	43;	DB 4;	
Matches	83.3%	Pred. No.	83;	
	5;	Mismatches	0;	Indels 0; Gaps 0;
QY	2	YWGGYW	7	
Db	55	HWGGYW	60	

RESULT 11

ID	ADN49403	standard; protein; 118 AA.		
XX	XX			
AC	ADN49403;			
XX	DT	01-JUL-2004 (first entry)		
XX	DE	Human anti-IL-8 variable heavy chain (VH) antibody homologue protein.		
XX	KW	Humanised antibody; therapy; diagnosis; autoimmune disease; cancer; transplant rejection; infectious disease; inflammation; variable heavy chain; VH; human.		
XX	OS	Homo sapiens.		
PN	US2004067532-A1.			
XX	PD	08-APR-2004.		
XX	PF	11-JUN-2003; 2003US-00460595.		
XX	PR	12-AUG-2002; 2002US-0403296P.		
XX	PA	(GENE-) GENETASTIX CORP.		
XX	PI	Zhu L, Wei S, Hua SB;		
XX	DR	WPI; 2004-355295/33.		
XX	PR	Screening library of humanized antibody useful in therapeutics, by expressing target protein in yeast expressing humanized antibodies, and selecting yeast in which reporter gene is expressed on activation by binding of antibody to protein.		
PS	DISCLOSURE;	SEQ ID NO 8; 38pp; English.		
XX	The present invention provides compositions, methods and kits for efficiently generating and screening humanised antibody with high affinity against a specific antigen. The invention is useful in therapeutics and the diagnosis of diseases such as autoimmune diseases, cancer, transplant rejection, infectious diseases and inflammation. The present sequence is human anti-IL-8 variable heavy chain (VH) antibody homologue protein.			
SQ	Sequence 118 AA;			
Query Match	77.8%	Score 42; DB 8;	Length 118;	
Best Local Similarity	83.3%	Pred. No. 1e+02;		
Matches	5;	Mismatches 0;	Indels 1;	Gaps 0;
QY	2	YWGGYW	7	
Db	101	YWGNYW	106	

ABU39959  
 ID ABU39959 standard; protein; 211 AA.  
 XX  
 AC ABU39959;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Pseudomonas putida.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-0094893.  
 PR 25-OCT-2001; 2001US-0342933P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 08-MAR-2002; 2002US-0362699P.  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselebeck R, Ohlsen KL, Zyskind JW,  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA,  
 XX  
 WPI; 2003-029926/02.  
 DR N-P5DB; ACA43829.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 Claim 25; SEQ ID NO 67883; 1766pp; English.  
 XX  
 The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation; or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC compound's activity; (8) manufacturing an antibiotic; (9) profiling a  
 CC product is overexpressed or underexpressed; (10) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)  
 XX  
 Sequence 211 AA;  
 SQ

Query Match 77.8%; Score 42; DB 6; Length 211;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Nismatches 1; Indels 0; Gaps 0;  
 PT - used in the treatment of leukaemia and B-cell lymphoma.  
 XX  
 Claim 2; Page 105-106; 14pp; English.

Qy 1 GYWGGYW 7  
 XX :|||  
 Db 36 GFWRGYW 42

RESULT 13  
 AAW09436  
 ID AAW09436 standard; protein; 243 AA.  
 XX  
 AC AAW09436;  
 XX DT 23-SEP-1997 (First entry)  
 DE Anti-CD19 antibody BLY3 single chain variable region.  
 XX KW CD19; antibody; variable region; B43; SJ25C1; Bly3; targeting;  
 KW cytotoxic conjugate; cancer treatment; leukaemia; B cell lymphoma;  
 KW variable; heavy; light chain.  
 XX OS Synthetic.  
 XX FH Key  
 FT Domain  
 FT Region  
 FT Peptide  
 FT Domain  
 FT Region  
 XX

Location/Qualifiers  
 1..114  
 /label= heavy\_chain  
 1..26  
 /label= FR1  
 27..31  
 /label= CDR1  
 32..45  
 /label= FR2  
 46..62  
 /label= CDR2  
 63..94  
 /label= FR3  
 95..114  
 /label= CDR3  
 115..129  
 /label= linker  
 110..243  
 /label= light\_chain  
 130..152  
 /label= FR1  
 153..167  
 /label= CDR1  
 168..182  
 /label= FR2  
 183..189  
 /label= CDR2  
 190..221  
 /label= FR3  
 222..243  
 /label= CDR3  
 PN WO9636336-A1.  
 XX  
 PD 21-NOV-1996.  
 XX  
 PF 15-MAY-1996; 95WO-US006941.  
 XX  
 PR 17-MAY-1995; 95US-00443408.  
 PA (MINU ) UNIV MINNESOTA.  
 PI Kersey JH, Bejcek BE, Wang D, Uckun FM;  
 XX DR WPI; 1997-042642/04.  
 DR N-PADB; AAT47734.  
 XX

CC AAW09436 represents a single chain variable region polypeptide (heavy chain-linker-light chain) of anti-cp19 antibody Bly3. Single chain polypeptides were conjugated to at least one cytotoxic agent and used to target the agent(s) to cancerous cells expressing CD19. The immunoconjugates can be used treatment of cancer, especially leukaemia and B-cell lymphoma  
 CC  
 XX  
 SQ Sequence 243 AA;  
 Query Match 77.8%; Score 42; DB 2; Length 243;  
 Best Local Similarity 83.3%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 PT -  
 XX  
 Claim 2; Page 105-106; 14pp; English.

Qy 2 YNGGYW 7  
 XX |||||  
 Db 97 YNGNYW 102

RESULT 14  
 AEM67619  
 ID AEM67619 standard; protein; 447 AA.  
 XX  
 AC AEM67619;  
 XX DT 20-NOV-2003 (First entry)  
 XX DE Photorhabdus luminescens protein sequence #716.  
 XX  
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; Plant; animal; microorganism; toxin;  
 KW antibiotic; bipesticide; virulence factor; disease model; plague;  
 KW whooping cough.  
 XX OS Photorhabdus luminescens.  
 XX PN WO200394867-A2.  
 XX  
 XX PD 28-NOV-2002.  
 XX PF 07-FEB-2002; 2002WO-1B003040.  
 XX  
 XX PR 07-FEB-2001; 2001FR-0001659.  
 XX PA (INSP ) INST PASTEUR.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Duchaud E, Tcaurit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 PI Buchrieser C;  
 XX DR 2003-148459/14.  
 XX  
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX  
 XX  
 PS Claim 2; SEQ ID NO 716; 1205pp; French.  
 XX  
 CC The invention relates to the isolation of genes and their encoded proteins from Photorhabdus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of P. luminescens and related species to study polymorphisms for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the proteins encoded by the genes are used for detection/identification of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than P. luminescens and are able to alter response or sensitivity to toxins and antibiotics produced by P. luminescens. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and

CC antibacterials useful as insecticides, bactericides and fungicides. The CC genes, proteins, vectors containing the genes and Ab are also useful. The CC therapeutic peptides (to treat microbial infection by bacteria or fungi that CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as CC biopesticides. Other uses of the genes and the proteins are as virulence CC factors and for identifying targets of human diseases for which P. CC luminescens is a model (particularly plague and whooping cough). This CC sequence represents one of the isolated P. luminescens proteins

XX Sequence 447 AA:

Query Match 77.8%; Score 42; DB 6; Length 447;  
Best Local Similarity 71.4%; Pred. No. 3.Se+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYRGGYW 7  
| : | || |  
Db 147 GFWNGYW 153

## RESULT 15

AAU90607 ID AAU90607 standard; peptide; 20 AA.  
XX AC AAU90607;  
XX DT 18-JUN-2002 (first entry)  
XX DE Insulin/insulin-like growth factor receptor-binding peptide #2563.  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.  
XX Synthetic.  
XX WO200172771-A2;  
XX PN 04-OCT-2001.  
XX PF 29-MAR-2000; 2000WO-US008528.  
XX PR 29-MAR-2000; 2000WO-US008528.  
XX PA (DGBTB-) DGI BIOTECHNOLOGIES LLC.  
PA (NOVO ) NOVO NORDISK AS.  
XX DR 2002-025774/03.

XX Modulating insulin activity in mammalian cells for treating e.g. diabetes and tumors, comprises using peptides that bind to insulin or insulin-like growth factor receptors.

XX PS Example 9; Fig 10A-2; 390pp; English.

XX The invention relates to a method of modulating insulin activity in CC mammalian cells by administering a peptide that binds the insulin receptor (IR). A composition containing a peptide, optionally expressed CC from gene therapy vectors, that binds to Site 1 of IR and an insulin CC agonist are useful for treating diabetes. Also, peptides that are CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are CC useful for treating insulin-like growth factor (IGF)-sensitive tumours CC (e.g. of prostate and breast) and diabetic neuropathy, while IGF-1 CC receptor agonists are useful for treating neurological diseases, CC including stroke and diabetic neuropathy. The peptides are also useful in CC screening for compounds that bind to IR or IGF-1 receptor, potential CC therapeutics and research reagents. AAU88034-AAU9057 represent IR and/or CC IGF-1 receptor-binding peptides and related amino acid sequences of the CC

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	440	15	US-10-282-122A-48322
2	48	88.9	296	14	US-10-156-761-9632
3	46	85.2	247	15	US-10-424-599-184877
4	45	83.3	141	16	US-10-437-963-143884
5	42	77.8	118	15	US-10-460-595-8
6	42	77.8	211	15	US-10-282-122A-67883
7	41	75.9	20	15	US-09-962-756-1075
8	41	75.9	20	15	US-10-925-471-1075
9	41	75.9	20	15	US-10-253-493-1075
10	41	75.9	26	10	US-09-962-756-1516
11	41	75.9	26	15	US-10-253-471-1516
12	41	75.9	26	15	US-10-253-493-1516
13	41	75.9	40	10	US-09-962-756-1113

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	440	15	US-10-282-122A-48322
2	48	88.9	296	14	Sequence 9632, AP
3	46	85.2	247	15	Sequence 184877, AP
4	45	83.3	141	16	Sequence 143884, AP
5	42	77.8	118	15	Sequence 8, AP
6	42	77.8	211	15	Sequence 67883, AP
7	41	75.9	20	15	Sequence 1075, AP
8	41	75.9	20	15	Sequence 1075, AP
9	41	75.9	20	15	Sequence 1075, AP
10	41	75.9	26	10	Sequence 1516, AP
11	41	75.9	26	15	Sequence 1516, AP
12	41	75.9	26	15	Sequence 1516, AP
13	41	75.9	40	10	Sequence 1113, AP

Sequence 1113, AP  
Sequence 1113, AP  
Sequence 246894,  
Sequence 745, AP  
Sequence 745, AP  
Sequence 745, AP  
Sequence 47856, AP  
Sequence 5061, AP  
Sequence 1129, AP  
Sequence 49552, AP  
Sequence 190291, AP  
Sequence 7823, AP  
Sequence 151557, AP  
Sequence 99, APPL  
Sequence 181262,  
Sequence 95, APPL  
Sequence 102, APPL  
Sequence 105, APPL  
Sequence 108, APPL  
Sequence 111, APPL  
Sequence 118, APPL  
Sequence 180291, AP  
Sequence 122848, AP  
Sequence 32, APPL  
Sequence 45092, AP  
Sequence 181854, AP  
Sequence 22275, AP  
Sequence 6849, AP  
Sequence 2, APPL  
Sequence 14, APPL

## ALIGNMENTS

RESULT 1  
US-10-282-122A-48322

; Sequence 48322, Application US/10282122A  
; Publication No. US2004029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cherry  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zybskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: EULTRA 034A  
; CURRENT APPLICATION NUMBER: US/10/282-122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 48322  
; LENGTH: 440  
; TYPE: PRT  
; ORGANISM: Bacteroides fragilis  
US-10-282-122A-48322

Query Match Score 54; DB 15; Length 440;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYGGGYW 7  
Db 212 GYGGYW 218

RESULT 2  
US-10-156-761-9632  
; Sequence 9632, Application US/10156761  
; Publication No. US20030119018A1  
GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIOUKI  
; APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 24-9-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-19  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 9632  
LENGTH: 296  
TYPE: PRT  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-9632

Query Match Score 48; DB 14; Length 296;  
Best Local Similarity 88.9%; Pred. No. 36%;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYGGGYW 7  
Db 198 GYNGAYW 204

RESULT 3  
US-10-424-599-184877  
; Sequence 184877, Application US/10424599  
; Publication No. US2004031072A1  
GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David J  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5322)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 143884  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4474C.1.dep

Query Match Score 45; DB 16; Length 141;  
Best Local Similarity 83.3%;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYGGYW 7  
Db 4 GCWGGINW 10

RESULT 4  
US-10-437-963-143884  
; Sequence 143884, Application US/10437963  
; Publication No. US20040123343A1  
GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5322)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 143884  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4474C.1.dep

Query Match Score 45; DB 16; Length 141;  
Best Local Similarity 83.3%;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYGGYW 7  
Db 109 GZWRGYW 115

RESULT 5  
US-10-460-595-8  
; Sequence 8, Application US/10460595  
; Publication No. US2004067532A1  
GENERAL INFORMATION:  
; APPLICANT: Zhi, Li  
; APPLICANT: Wei, Shuanghong  
; APPLICANT: Hua, Shaobing  
; APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5322)B  
CURRENT APPLICATION NUMBER: US/10/460,595  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: US 60/403,296  
PRIOR FILING DATE: 2002-08-12  
NUMBER OF SEQ ID NOS: 9

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-460-595-8

Query Match      77.8%; Score 42; DB 15; Length 118;
Best Local Similarity 83.3%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 1;

Qy   2 YWGTYW 7
Db   101 YWGNYW 106

RESULT 6
US-10-282-122A-67883
; Sequence 67883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA-03A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-04-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NCS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 67883
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-67883

Query Match      77.8%; Score 42; DB 15; Length 211;
Best Local Similarity 71.4%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1;

Qy   1 GYWGTYW 7
Db   36 GFWRGYW 42

```

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0; OTHER INFORMATION: peptide
Qy   3 WGGYW 7 ; FEATURE: MOD RES
Db   13 WGGYW 17 ; LOCATION: (19)
; OTHER INFORMATION: Unknown amino acid
US-09-562-756-1516

Query Match 75.9%; Score 41; DB 10; Length 26;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy   1 GYWGTYW 7
Db   20 GFWGGSW 26

RESULT 11
US-10-253-471-1516
; Sequence 1516, Application US/10253471
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-14
; PRIOR APPLICATION NUMBER: 09/362,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1075
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-493-1075

Query Match 75.9%; Score 41; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy   3 WGGYW 7
Db   13 WGGYW 17

RESULT 10
US-09-962-756-1516
; Sequence 1516, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLOME, ARTHUR J.
; APPLICANT: SCHAFFER, LAUKE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-1516
; Sequence 1516, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-1516
; Sequence 1516, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (19)
; OTHER INFORMATION: Unknown amino acid
US-10-253-493-1516

Query Match    75.9%; Score 41; DB 15; Length 26;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 WGGYW 7
Db     20 GFWGGSW 26

RESULT 13
; Sequence 1113, Application US/09962756
; Publication No. US2003019514A1

; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; APPLICANT: BRISSETTE, RENE
; APPLICANT: BLOMB, ARTHUR J.
; APPLICANT: SCHAFER, LAUKE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; PRIOR APPLICATION NUMBER: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR APPLICATION NUMBER: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1113
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-09-962-756-1113

Query Match    75.9%; Score 41; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 WGGYW 7
Db     33 WGGYW 37

RESULT 15
; Sequence 1113, Application US/10253493
; Publication No. US20040023887A1

; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; APPLICANT: BRISSETTE, RENE
; APPLICANT: BLOMB, ARTHUR J.
; APPLICANT: SCHAFER, LAUKE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-40561US1
; CURRENT APPLICATION NUMBER: US/10/253,493
; PRIOR APPLICATION NUMBER: 2002-09-24
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1113
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-10-253-493-1113

Query Match    75.9%; Score 41; DB 15; Length 40;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 WGGYW 7
Db     33 WGGYW 37

RESULT 14
; Sequence 1113, Application US/10253471
; Publication No. US20030236190A1
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-10-253-471-1113

```

Tue Apr 19 06:58:37 2005

seq1.rapb

Page 6

Db 33 WGGYW 37

Search completed: April 18, 2005, 20:29:25  
Job time : 52.2791 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
1	100.0	339	4	US-09-252-991A-26841	Sequence 26841, A		
2	83.3	500	4	US-09-252-991A-21214	Sequence 21214, A		
3	75.9	105	4	US-09-205-258-745	Sequence 745, App		
4	75.9	571	4	US-09-902-540-14983	Sequence 14983, A		
5	74.1	585	1	US-08-212-188-4	Sequence 4, App1		
6	74.1	585	3	PCT-US95-02708-4	Sequence 4, App1		
7	74.1	585	5	PCT-US95-02708-4	Sequence 4, App1		
8	74.1	1827	4	US-09-443-780C-14	Sequence 14, App1		
9	74.1	1827	4	US-09-079-727-3	Sequence 3, App1		
10	72.2	1621	3	US-08-972-927-3	Sequence 6, App1		
11	72.2	1622	3	US-08-972-927-6	Sequence 6, App1		
12	70.4	704	12	US-08-241-054-99	Sequence 99, App1		
13	70.4	12	1	US-08-439-817-99	Sequence 99, App1		
14	70.4	12	1	US-08-485-108-99	Sequence 99, App1		
15	70.4	86	4	US-09-248-796A-22524	Sequence 22524, A		
16	70.4	212	4	US-09-543-681A-5065	Sequence 5065, App		
17	70.4	470	4	US-09-252-991A-23310	Sequence 23310, A		
18	70.4	638	2	US-08-846-762-95	Sequence 95, App1		
19	68.5	80	4	US-09-252-991A-22806	Sequence 22806, A		
20	68.5	150	4	US-09-371-056-16	Sequence 16, App1		
21	68.5	154	4	US-09-902-540-16787	Sequence 16787, A		
22	68.5	231	3	US-08-448-489-19	Sequence 19, App1		
23	68.5	231	4	US-09-252-991A-23310	Sequence 19, App1		
24	68.5	275	4	US-09-371-056-10	Sequence 10, App1		
25	68.5	355	1	US-08-118-270-11	Sequence 11, App1		
26	68.5	355	1	PCT-US93-08528-11	Sequence 11, App1		
27	68.5	489	1	US-08-489-733-6	Sequence 6, App1		

## ALIGNMENTS

RESULT 1  
US-09-252-991A-26841  
; Sequence 26841, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26841

Query Match 100.0% Score 54; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 1.4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYWGTY 7  
Db 261 GYWGTY 267

RESULT 2  
US-09-252-991A-21214  
; Sequence 21214, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO. 21214  
; LENGTH: 500  
; TYPE: PRT

ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-21214

Query Match Score 45; DB 4; Length 500;  
Best Local Similarity 83.3%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYWGTYW 7  
Db 331 GYWGRW 337

RESULT 3  
US-09-205-258-745  
Sequence 745 Application US/09205258

Patent No. 6525174  
GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

CURRENT APPLICATION NUMBER: US/09/205,258

CURRENT FILING DATE: 1998-12-04

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-05

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-05

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-05

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-05

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,974  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,917  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,918  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,893  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,920  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,878  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,921  
; Earlier Filing Date: 1997-06-06 ; Earlier Application Number: 60/048,922  
; Earlier Filing Date: 1998-07-15 ; Earlier Application Number: 60/094,657  
; Earlier Filing Date: 1998-07-30 ; Number of SEQ ID NOS: 1227  
; Software: Patentin Ver. 2.0 ; SEQ ID NO: 745  
; LENGTH: 105 ; TYPE: PRT ; ORGANISM: Homo sapiens  
; US-09-205-258-745

Query Match Score 41; DB 4; Length 105;  
Best Local Similarity 75.9%; Pred. No. 33; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYWGTYW 7  
Db 11 GYWGRW 17

RESULT 4  
US-09-902-540-14983  
; Sequence 14983, Application US/09902540  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: *Myxococcus xanthus Genome Sequences and Uses Thereof*  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2004-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO: 14983  
; LENGTH: 571  
; TYPE: PRT  
; ORGANISM: *Myxococcus xanthus*  
; US-09-902-540-14983

Query Match Score 41; DB 4; Length 571;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 WGGYW 7

```

ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/970,725
FILING DATE: 14-NOV-1997
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/212,188
FILING DATE: 16-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weisser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE DOCKET NUMBER: 372,6601P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-970-725-4

Query Match Score 40; DB 3; Length 585;
Best Local Similarity 93.3%; Pred. No. 2,3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YNGGYW 7
Db 109 YNGRYW 114

RESULT 7
PCT-US95-02708-4
SEQUENCE 4, Application PC/TUS9502708
GENERAL INFORMATION:
APPLICANT: BECKER, JEFFREY M.
APPLICANT: BECKER, JEFFREY M.
TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
NUMBER OF SEQUENCES: 11
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-970-725-4

Query Match Score 40; DB 3; Length 585;
Best Local Similarity 93.3%; Pred. No. 2,3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YNGGYW 7
Db 109 YNGRYW 114

RESULT 6
US-08-970-725-4
SEQUENCE 4, Application US/08970725
PATENT NO. 6080542
GENERAL INFORMATION:
APPLICANT: Becker, Jeffrey M.
APPLICANT: Stacey, Gary
TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
NUMBER OF SEQUENCES: 14
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
CITY: Philadelphia
STATE: PA
COUNTRY: USA

Query Match Score 40; DB 1; Length 585;
Best Local Similarity 83.3%; Pred. No. 2,3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YNGGYW 7
Db 109 YNGRYW 114

RESULT 5
US-08-970-725-4
SEQUENCE 4, Application US/08212188
PATENT NO. 5689039
GENERAL INFORMATION:
APPLICANT: BECKER, JEFFREY M.
APPLICANT: STACEY, GARY M.
TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
NUMBER OF SEQUENCES: 11
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,188
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AUBERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE DOCKET NUMBER: 7493-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: ATPTR23p
US-08-212-188-4

Query Match Score 40; DB 1; Length 585;
Best Local Similarity 83.3%; Pred. No. 2,3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YNGGYW 7
Db 109 YNGRYW 114

```

MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Arabidopsis thaliana  
 IMMEDIATE SOURCE:  
 CLONE: ATPTR2Bp  
 PCT-US95-02708-4

Query Match 74.1%; Score 40; DB 5; Length 585;  
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGYW 7  
 DB 109 YWGRYW 114

RESULT 8  
 US-09-443-780C-14  
 Sequence 14, Application US/09443780C  
 Patent No. 6699973

GENERAL INFORMATION:  
 APPLICANT: O'Mahony, Daniel J.  
 APPLICANT: Seveso, Michaela  
 TITLE OF INVENTION: Antibodies To Peptides That Target GIT Receptors And Related Methods  
 FILE REFERENCE: E106/2003/07  
 CURRENT APPLICATION NUMBER: US/09/443,780C  
 CURRENT FILING DATE: 1999-11-19  
 PRIOR APPLICATION NUMBER: US 60/109,036  
 PRIOR FILING DATE: 1998-11-19  
 NUMBER OF SEQ ID NOS: 84  
 SEQ ID NO: 14  
 LENGTH: 1827  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: hSI receptor  
 US-09-443-780C-14

Query Match 74.1%; Score 40; DB 4; Length 1827;  
 Best Local Similarity 71.4%; Pred. No. 6.5e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGGYW 7  
 DB 1491 GRWGSHW 1497

RESULT 9  
 US-09-079-723-179  
 Sequence 179, Application US/09079723  
 Patent No. 67031362

GENERAL INFORMATION:  
 APPLICANT: Alvarez, Vernon L.  
 APPLICANT: O'Mahony, Daniel J.  
 APPLICANT: Lambkin, Imelda J.  
 APPLICANT: Singletton, Judith  
 APPLICANT: Patterson, Catherine A.  
 APPLICANT: Cagney, Gerard M.  
 APPLICANT: Belinka, Benjamin A.  
 APPLICANT: Carter, John M.  
 TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Penne & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/079,723  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Misrock, S. Leslie  
 REGISTRATION NUMBER: 18-872  
 REFERENCE/DOCKET NUMBER: 1101-219  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-669-9741  
 TELEX: 65141 PENNIE  
 INFORMATION FOR SEQ ID NO: 179:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1827 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 US-09-079-723-179

Query Match 74.1%; Score 40; DB 4; Length 1827;  
 Best Local Similarity 71.4%; Pred. No. 6.5e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGGYW 7  
 DB 1491 GRWGSHW 1497

RESULT 10  
 US-08-972-927-3  
 Sequence 3, Application US/08972927  
 Patent No. 6166290

GENERAL INFORMATION:  
 APPLICANT: Rea, Philip A.  
 APPLICANT: Lu, Yu-Ping  
 APPLICANT: Li, Ze-Sheng  
 TITLE OF INVENTION: GLUTATHIONE-S-CONUGATE TRANSPORT IN PLANTS  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PANITCH SCHWARZE JACOBS & NADELL, P.C.  
 STREET: One Commerce Square, 2005 Market Street, 22nd Floor  
 CITY: Philadelphia  
 STATE: Pennsylvania  
 COUNTRY: US  
 ZIP: 19103-7086

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/972,927  
 FILING DATE: 18-NOV-1997  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/031,040  
 FILING DATE: 18-NOV-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/061,328  
 FILING DATE: 08-OCT-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Doyle Leary Ph.D., Kathryn  
 REGISTRATION NUMBER: 36,317  
 REFERENCE/DOCKET NUMBER: 9596-12U2  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-965-1284  
 TELEFAX: 215-567-2991  
 TELEX: 831-494  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1621 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-972-927-3

Query Match 72.2%; Score 39; DB 3; Length 1621;  
 Best Local Similarity 66.7%; Pred. No. 8.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YWGGYW 7  
 Db 303 WWGGFW 308

RESULT 11  
 US-08-972-927-6  
 Sequence 6, Application US/08972927  
 Patent No. 6166390

GENERAL INFORMATION:  
 APPLICANT: Rea, Philip A  
 APPLICANT: Lu, Yu-Ping  
 APPLICANT: Li, Ze-Sheng  
 TITLE OF INVENTION: GLUTATHIONE-S-CONJUGATE TRANSPORT IN  
 TITLE OF INVENTION: PLANTS  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PANTICH SCHWARZE JACOBS & NADBL, P.C.  
 STREET: One Commerce Square, 2005 Market Street, 22nd  
 STREET: Floor  
 CITY: Philadelphia  
 STATE: Pennsylvania  
 COUNTRY: US  
 ZIP: 19103-7086

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/972, 927  
 FILING DATE: 18-NOV-1997  
 CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/031, 040  
 FILING DATE: 18-NOV-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/061, 328  
 FILING DATE: 08-OCT-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Doyle Leary Ph.D.; Kathryn  
 REGISTRATION NUMBER: 36,317  
 REFERENCE/DOCKET NUMBER: 9596-12U2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-965-1284  
 TELEX: 831-494  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1622 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-972-927-6

Query Match 72.2%; Score 39; DB 3; Length 1622;

Best Local Similarity 66.7%; Pred. No. 8.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 YWGGYW 7  
 Db 303 WWGGFW 308

RESULT 12  
 US-08-241-054-99  
 Sequence 99, Application US/08241054  
 Patent No. 5643873

GENERAL INFORMATION:  
 APPLICANT: Barrett, Ronald W.  
 APPLICANT: Cwirka, Steven E.  
 APPLICANT: Dower, William J.  
 APPLICANT: Koller, Kerry J.  
 APPLICANT: Lee, Jung  
 APPLICANT: Martens, Beatrice L.  
 APPLICANT: Ruhland-Fritsch, Beatrice  
 TITLE OF INVENTION: Peptides and Compounds That Bind Leukocyte Adhesion  
 TITLE OF INVENTION: Selectins Including Endothelium Molecule 1  
 NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: 699 Prince Street  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22313

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/241, 054  
 FILING DATE: 11-MAY-1994  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/057, 295  
 FILING DATE: 05-MAY-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/881, 395  
 FILING DATE: 06-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gerald F. Swiss  
 REGISTRATION NUMBER: 30,113  
 REFERENCE/DOCKET NUMBER: 000324-002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-854-7400  
 TELEX: 415-854-8275  
 INFORMATION FOR SEQ ID NO: 99:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 12 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-08-241-054-99

Query Match 70.4%; Score 38; DB 1; Length 12;  
 Best Local Similarity 71.4%; Pred. No. 12;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GWGGYW 7  
 Db 2 GWGGLW 8

RESULT 13  
 US-08-439-817-79

Sequence 79, Application US/08439817  
 GENERAL INFORMATION:  
 APPLICANT: Barrett, Ronald W.  
 APPLICANT: Cwirla, Steven E.  
 APPLICANT: Dower, William J.  
 APPLICANT: Koller, Kerry J.  
 APPLICANT: Lee, Jung  
 APPLICANT: Martens, Christine L.  
 APPLICANT: Ruhland-Fritsch, Beatrice  
 TITLE OF INVENTION: Peptides and Compounds That Bind Leukocyte Adhesion  
 TITLE OF INVENTION: Selectins Including Endothelial Leukocyte Adhesion  
 NUMBER OF SEQUENCES: 162  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Affymax Technologies, NV  
 STREET: 4001 Miranda Ave.  
 CITY: Palo Alto  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/439,817  
 FILING DATE: 12-MAY-1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/241,054  
 FILING DATE: 11-MAY-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/439,817  
 FILING DATE: 12-MAY-1995  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/241,054  
 FILING DATE: 11-MAY-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens, Lauren L.  
 FILING DATE: 05-MAY-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/881,395  
 FILING DATE: 06-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens, Lauren L.  
 REGISTRATION NUMBER: 36,691  
 REFERENCE/DOCKET NUMBER: 000324-002/1056  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-496-2300  
 TELEFAX: 415-424-0832  
 INFORMATION FOR SEQ ID NO: 79:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 12 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-439-817-79

Query Match 1 GYWGQYW 7  
 Best Local Similarity 71.4%; Pred. No. 12; Length 12;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 1 GYWGQYW 7  
 Best Local Similarity 70.4%; Score 38; DB 1; Length 12;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 15  
 US-09-248-796A-22524  
 Sequence 22524, Application US/09248796A  
 GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al  
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 FILE REFERENCE: 107196\_132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13

RESULT 14  
 US-08-485-508-99  
 Sequence 99, Application US/08485508  
 GENERAL INFORMATION:  
 APPLICANT: Barrett, Ronald W.  
 APPLICANT: Cwirla, Steven E.  
 APPLICANT: Dower, William J.  
 APPLICANT: Koller, Kerry J.

```
; NUMBER OF SEQ ID NOS : 28208
; SEQ ID NO 22524
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22524

Query Match          Score 38; DB 4; Length 86;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          2 YWGGYW 7
           :||| |
Db          10 FWGNYW 15
```

Search completed: April 18, 2005, 20:00:37  
Job time : 28.5349 secs

## *Form 1 Rudge Blank (uspto)*

Copyright GenCore version 5.1.6  
(c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 06:18:13 ; Search time 44 Seconds  
(without alignments)  
19.681 Million cell updates/sec

Title: SEQ2  
Perfect score: 60  
Sequence: 1 xgyngkgyw 9

Scoring table: BLOSUM62DX  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

Database : PIR\_79;\*  
1: Piri1;\*  
2: Piri2;\*  
3: Piri3;\*  
4: Piri4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	80.0	1365	1 BVBYK5	killer toxin resis
2	45	75.0	441	2 E6994	short-chain fatty
3	44.5	74.2	83	2 E69903	hypothetical prote
4	44.5	74.2	187	2 G83047	hypothetical prote
5	44	73.3	358	2 H83554	hypothetical prote
6	43	71.7	80	2 AC2394	hypothetical prote
7	43	71.7	180	2 D83812	hypothetical prote
8	43	71.7	470	2 B75567	hypothetical prote
9	43	71.7	1311	2 A56390	mannosyl-glycoprot
10	43	71.7	1312	2 E95006	beta-N-acetylhexos
11	43	71.7	1312	2 A97879	hypothetical prote
12	42	70.0	181	2 A83915	hypothetical prote
13	42	70.0	196	2 AB2918	acetyltransferase
14	42	70.0	207	2 H97692	conserved hypothet
15	42	70.0	375	2 AI2041	hypothetical prote
16	42	70.0	563	2 A69491	probable formylmet
17	42	70.0	844	2 I37079	DNA ligase (NTP)
18	41	68.3	113	2 PH1663	Ig heavy chain V r
19	41	68.3	145	2 F75337	transposase - Dein
20	41	68.3	1447	2 S63669	doglucose-glycopr
21	40	66.7	102	2 B72742	hypothetical prote
22	40	66.7	194	2 H2845	hypothetical prote
23	40	66.7	228	2 S26009	hypothetical prote
24	40	66.7	250	1 S70854	probable heme tran
25	40	66.7	489	2 S77357	hypothetical prote
26	40	66.7	495	2 AD1927	probable ABC trans
27	40	66.7	554	1 A28716	probable ABC trans
28	40	66.7	589	2 A29476	probable ABC trans
29	40	66.7	589	2 B29514	probable ABC trans

## ALIGNMENTS

RESULT 1

BVBYK5 killer toxin resistance protein KRE5 precursor - yeast (Saccharomyces cerevisiae)

N: Alternative names: protein Q2254; protein YOR336w

C: Species: Saccharomyces cerevisiae

C: Accession: S62066; A36327; S67243; S71974; S12202

R: Parle, A.G.; Hand, N.J.; Goulding, S.G.; Wolfe, K.H.

submitted to the EMBL Data Library, June 1995

A: Description: Sequence of 29 kilobases around the PDR10 locus on the right arm of Sacc.

A: Reference number: A36327; MUID: 90258892; PMID: 2188106

A: Accession: A36327

A: Molecule type: DNA

A: Residues: 1-11365 <PAR>

A: Cross-references: UNIPROT:P22023; EMBL:749821; PID:CAA9981.1; PID:91163062; PID:CAA9981.1; PID:9171795

A: Cross-references: EMBL:133556; PID:9171794; PID:AAA34125.1; PID:9171795

R: Meaden, P.; Hill, K.; Wagner, J.; Slipetz, D.; Sommer, S.S.; Bussey, H.

Mol. Cell. Biol. 10, 3013-3019, 1990

A: Title: The yeast KRE5 gene encodes a probable endoplasmic reticulum protein required for normal cell growth

A: Reference number: A36327; MUID: 90258892; PMID: 2188106

A: Accession: A36327

A: Molecule type: DNA

A: Residues: 1-581,583-779,1,'781,'1'KMKCOKONISK,'794,'K,'795-1365 <MEA>

A: Cross-references: EMBL:1-581,583-779,1,'781,'1'KMKCOKONISK,'794,'K,'795-1365 <MEA>

R: Goulding, S.E.; Hand, N.J.; Parle-McDermott, A.G.; Wolfe, K.H.

submitted to the Protein Sequence Database, July 1996

A: Reference number: S67233

A: Accession: S67243

A: Molecule type: DNA

A: Residues: 1-1365 <GOI>

A: Cross-references: EMBL:275244; PID:91420730; PID:CAA9959.1; PID:91420731; GSPDB:GN0

A: Experimental source: strain S288C

R: Parle-McDermott, A.G.; Hand, N.J.; Goulding, S.E.; Wolfe, K.H.

Yeast 12, 999-1004, 1996

A: Title: Sequence of 29 kb around the PDR10 locus on the right arm of Saccharomyces cer

A: Reference number: S71966; MUID: 97051586; PMID: 8896263

A: Accession: S71974

A: Molecule type: DNA

A: Residues: 1-1365 <PAR>

A: Cross-references: ENBL:Z49821; PID:91163062; PID:CAA89981.1; PID:91163070

A: Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C: Genetics:

A: Gene: SGD:KRE5; MIPS: YOR336w

A: Cross-references: SGD:YOR336w

A: Map position: 15R

C: Function:

A: Description: required for normal cell growth

A: Pathway: (1->6)-beta-D-glucan biosynthesis

C: Superfamily: KRE5 protein

C: Keywords: endoplasmic reticulum; glycoprotein

F: 1-17/Domain: signal sequence #status predicted <SIG>

F: 18-1365/Product: killer toxin resistance protein KRE5 #status predicted <MAT>

F;11362-1365/Region: endoplasmic reticulum retention signal #status predicted  
F;1115,-228,193,457,519,523,604,644,870,1091,1150,1195/Binding site: carbohydrate (Asn)  
C;Species: Bacillus halodurans (strain C-125)  
C;Accession: F83694  
C;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A;Reference number: A83650; MUID:11058132  
A;Accession: F83694  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q9KFM3; GB:AP001508; PIDN:BA000004; NID:g1012890; PIDN:BA00400  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: atoE  
C;Superfamily: conserved hypothetical integral membrane protein HP0693  
Query Match Score 45; DB 2; Length 44;  
Best Local Similarity 75.0%; Pred. No. 20;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 YWKGKYW 9  
Db 48 YWGEGYW 54

RESULT 4  
GB:047  
hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: G83047  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.I.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: G83047  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q9HV15; GB:AE004892; PIDN:AA004091; NID:g9951049; PIDN:AA0043  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA4793

Query Match Score 74.2%; DB 2; Length 187;  
Best Local Similarity 77.8%; Pred. No. 10;  
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 XGYWGKGYW 9  
Db 108 GGYWG-GYW 115

RESULT 5  
H83554  
hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: H83554  
R;Stover, C.K.; Pham, X.Q.; Brody, L.I.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: H83554  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q9T5J3; GB:AE004508; PIDN:AA004091; NID:g9946611; PIDN:AA00415  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA0736

Query Match Score 73.3%; DB 2; Length 358;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 XGYWGKGYW 9  
Db 197 VGTWGGGYW 205

RESULT 6  
AC2394  
hypothetical protein asr4707 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

Query Match Score 44.5; DB 2; Length 83;  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residue 1-83 <UNK>  
A;Cross-references: UNIPROT:O34654; GB:Z99114; PIDN:G2634230; NID:g13852.  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yodI

C;Date: 14-Dec-2001 # sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C;Accession: AC2394  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, S.; Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AC2394  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-80 <STO>  
 A;Cross-references: UNIPROT:Q8YN63; GB:BA000019; PIDN:BAB76406.1; PID:917133844; GSPDB:G  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: asr4707

Query Match 71.7%; Score 43; DB 2; Length 80;  
 Best Local Similarity 75.0%; Pred. No. 7.7;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XGYWKGKY 8  
 :||| |  
 Db 37 GGYWGSKY 44

RESULT 7  
 D83812 hypothetical protein BH1300 [imported] - Bacillus halodurans (strain C-125)  
 C;Species: Bacillus halodurans  
 C;Date: 01-Dec-2000 # sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C;Accession: D83812  
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A;Reference number: A83650; MUID:20512582; PMID:11058132  
 A;Accession: D83812  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-180 <STO>  
 A;Cross-references: UNIPROT:9RKDB3; GB:AP001511; GB:BA000004; PID:G1017727; PIDN:BAB050  
 C;Genetics:  
 A;Gene: BH1300

Query Match 71.7%; Score 43; DB 2; Length 180;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XGYWKGKY 8  
 :||| |  
 Db 100 QSYWKGKY 107

RESULT 8  
 B72567 hypothetical protein APR1820 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Accession: B72567  
 R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudooh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999  
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A;Reference number: A72450; MUID:99310339; PMID:10382966  
 A;Accession: B72567  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-470 <STO>  
 A;Cross-references: UNIPROT:Q9YAX5; DDBJ:AP000062; NID:95105244; PIDN:BAA80823.1; PID:95  
 C;Genetics:  
 A;Gene: AP1820  
 C;Superfamily: phytene dehydrogenase

RESULT 9  
 A56390 mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - Streptococcus pneumoniae (strain TIGR4)  
 C;Species: Streptococcus pneumoniae  
 C;Accession: A56390  
 R;Clarke, V.A.; Platt, N.; Butters, T.D.  
 J. Biol. Chem. 270, 8805-8814, 1995  
 A;Title: Cloning and expression of the beta-N-acetylglucosaminidase gene from Streptococcus pneumoniae  
 A;Reference number: A56390; MUID:95238375; PMID:7721787  
 A;Accession: A56390  
 A;Status: preliminary  
 A;Cross-references: UNIPROT:P49610; GB:L36923; NID:9784896; PIDN:AC041450.1; PID:9784899  
 C;Genetics:  
 A;Gene: strH  
 C;Keywords: glycosidase; hydrolase; tandem repeat  
 F;1-33/Domain: signal sequence #status predicted <STG>

Query Match 71.7%; Score 43; DB 2; Length 1311;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKY 9  
 :||| |  
 Db 868 ISWWSKGWW 876

RESULT 10  
 E95006 beta-N-acetylhexosaminidase [imported] - Streptococcus pneumoniae (strain TIGR4)  
 C;Species: Streptococcus pneumoniae  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 C;Accession: E95006  
 R;Petterson, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei  
 nson, T.; Hickey, B.K.; Holt, I.B.  
 J.D. Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
 Science 293, 498-506, 2001  
 A;Authors: Loftus, B.J.; Yang, P.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison  
 A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A;Reference number: A95000; MUID:21357209; PMID:11463916  
 A;Accession: E95006  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1312 <KUR>  
 A;Experimental source: strain TIGR4  
 C;Genetics:  
 A;Gene: SP0057

Query Match 71.7%; Score 43; DB 2; Length 1312;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKY 9  
 :||| |  
 Db 869 ISWWSKGWW 877

RESULT 11  
 A97879 A;Cross-references: UNIPROT:Q9YAX5 [imported] - Streptococcus pneumoniae (strain

C;Species: Streptococcus pneumoniae  
 C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C;Accession: A97879  
 R;Hockins,J.A.;Alborn Jr., W.;Arnold, J.;Bleszczak, L.;Burgett, S.;DeHoff, B.S.;E  
 e, R.;LeBlanc, D.J.;Lee, L.N.;Lefkowitz, E.J.;Lu, J.;Matsushima, P.;McArehn, S.;M  
 Y, P.;Sun, P.M.;Winkler, M.E.  
 J;Bacteriol. 183, 5709-5717, 2001  
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A;Reference number: A97872; MUID:2142945; PMID:11544234  
 A;Status: preliminary  
 A;Residues: 1-1312 <KUR>  
 A;Cross-references: UNIPROT:Q8DR16; GB:AE007317; PIDN:ARK98861.1; PID:915457590; GSPDB:C  
 A;Gene: strH  
 C;Keywords: glycosidase; hydrolase

Query Match 71.6%; Score 43; DB 2; Length 1312;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XYGKGYW 9  
 Db 869 ISWNSKGWW 877

RESULT 12  
 A83915  
 C;Species: Bacillus halodurans (strain C-125)  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C;Accession: A83915  
 R;Takami, H.; Nakagawa, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A;Reference number: A83650; MUID:20512582; PMID:11058132  
 A;Accession: A83915  
 A;Status: preliminary  
 A;Residues: 1-181 <STO>  
 A;Cross-references: UNIPROT:Q9KB15; GB:APP001514; GB:BA000004; NID:910174613; PIDN:BAB058  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: BH2121  
 C;Superfamily: ribosomal-protein-serine N-acetyltransferase

Query Match 70.0%; Score 42; DB 2; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YWGKGY 8  
 Db 100 YWGKGY 105

RESULT 13  
 AE2918  
 C;Species: Agrobacterium tumefaciens  
 C;Accession: AE2918  
 R;Hood, D.W.; Setubal, J.C.; Kaul, R.; Marks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
 erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 i; Karp, P.; Romero, P.; Zhang, S.  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AE2918  
 A;Status: preliminary

A;Molecule type: DNA  
 A;Residues: 1-196 <KUR>  
 A;Cross-references: UNIPROT:Q8UBR8; GB:AB008688; PIDN:AAJ43763.1; PID:917741299; GSPDB:C  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atu2782  
 A;Map position: circular chromosome

Query Match 70.0%; Score 42; DB 2; Length 196;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YWGKGY 8  
 Db 116 YWGKGY 121

RESULT 14  
 H97692  
 conserved hypothetical protein BH0837 (AP001510) [imported] - Agrobacterium tumefaciens  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C;Accession: H97692  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: H97692  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-207 <KUR>  
 A;Cross-references: UNIPROT:Q8UBR8; GB:AB007869; PIDN:ANK88497.1; PID:915158008; GSPDB:C  
 C;Genetics:  
 A;Gene: AGR\_C\_5049  
 A;Map position: circular chromosome

Query Match 70.0%; Score 42; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YWGKGY 8  
 Db 127 YWGKGY 132

RESULT 15  
 A12041  
 hypothetical protein all11887 [imported] - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp. PCC 7120  
 C;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C;Accession: A12041  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuch  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, t  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: A12041  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-375 <KUR>  
 A;Cross-references: UNIPROT:08YVT7; GB:BA000019; PIDN:BAB73586.1; PID:917130977; GSPDB:C  
 A;Genetics: all11887

Query Match 70.0%; Score 42; DB 2; Length 375;  
 Best Local Similarity 55.6%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWGKGY 9  
 :||| |

Tue Apr 19 06:58:39 2005

seq2.rpr

Page 5

Db 95 NGFWNGGW 103

Search completed: April 19, 2005, 06:43:17  
Job time : 46 secs

This Page Blank (uspto)

Scoring table:	BLOSUM6DX	ALIGMENTS			
Gapop:	10.0 , Gapext 0.5				
Searched:	1612378 seqs, 512079187 residues				
Total number of hits satisfying chosen parameters:	1612378				
Minimum DB seq length:	0				
Maximum DB seq length:	2000000000				
Post-processing:	Minimum Match 0% Maximum Match 100%				
Database :	UniProt_03: 1: uniprot_sprot: 2: uniprot_trembl: * * uniprot_trembl: *				
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES					
Result No.	Query	Match	Length	DB ID	Description
1	50	83.3	840	2 Q8A6N9	Q8a6n9 bacterooides
2	50	83.3	640QS3	Q6dgs3 bacterooides	
3	49	81.7	214	2 Q8ABU4	Q8au4 bacterooides
4	49	81.7	841	2 Q7MUI8	Q7mu18 porphromon
5	48	80.0	210	2 Q8AH510	Q8s51 bacteroides
6	48	80.0	748	2 Q9HLG7	Q9hl97 thermoplasm
7	48	80.0	1293	2 Q759B3	Q759b3 ashbya goss
8	48	80.0	1365	1 KRB5_YEAST	P22023 achromyc
9	46	76.7	574	1 ZBT3_HUMAN	Q9hsj0 homo sapien
10	45.5	75.8	74	2 Q63VFT7	Q63vf7 burkholderi
11	45	75.0	198	2 Q9AXA8	Q9axa8 oryza sativa
12	45	75.0	441	2 Q9KFW3	Q9ktw3 bacillus ha
13	45	75.0	1326	2 Q6FV20	Q6tv20 candida gla
14	44.5	74.2	83	1 YODI_BACSU	Q34654 bacillus su
15	44.5	74.2	187	2 Q9HV15	Q9av15 pseudomonas
16	44.5	74.2	189	2 Q8QB8	Q8qgb8 pseudomonas
17	44.5	74.2	440	2 Q64YU6	Q64yr6 bacterooides
18	44	73.3	120	2 Q8BSY2	Q8bsy2 diachasmimo
19	44	73.3	358	2 Q9IJ5J3	Q9ij5j3 pseudomonas
20	44	73.3	443	2 Q8EH19	Q8ah39 shewarella
21	44	73.3	716	2 Q9HLF3	Q9hf3 thermoplasm
22	43.5	72.5	181	2 Q63BF4	Q63bf4 bacillus ce
23	43.5	72.5	182	2 Q734HO	Q734ho bacillus ce
24	43	71.7	80	2 Q8YN63	Q8yn63 anabaena sp
25	43	71.7	127	2 Q858P1	Q8s8p1 lactococcus
26	43	71.7	133	2 Q21906	Q21906 bacteriophila
27	43	71.7	180	2 Q9KDB3	Q9kb3 bacillus ha
28	43	71.7	277	2 Q9F3T0	Q9f3t0 rhodothermu
29	43	71.7	314	2 Q8A188	Q8a188 bacterooides
30	43	71.7	446	2 Q7VX54	Q7vx54 bordetella
31	43	71.7	468	2 Q7WLGS	Q7wlgs bordetella

Run on:	April 19, 2005, 05:59:25 ; Search time 176 Seconds (without alignments) 26.186 Million cell updates/sec	RESULT 1
Title:	SEQ2	Q8A6N9 PRELIMINARY; PRT; 840 AA.
Perfect score:	60	ID Q8A6N9; AC Q8A6N9; DT 01-JUN-2003 (TrEMBLrel. 24, Created) DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update) DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
Sequence:	1 xgywgkgyw 9	DE Putative alanyl dipeptidyl peptidase.
Scoring table:	BLOSUM6DX	GN Ordered locust names-BTR1838;
Gapop:	10.0 , Gapext 0.5	OS Bacteroides thetaiotaomicron.
Searched:	1612378 seqs, 512079187 residues	OC Bacteroidaceae; Bacteroides (class); Bacteroidales; OX NCF-TaxID=818;
Total number of hits satisfying chosen parameters:	1612378	RN
Minimum DB seq length:	0	RP SEQUENCE FROM N.A.
Maximum DB seq length:	2000000000	RC STRAIN=VP-5492 / ATCC 29148;
Post-processing:	Minimum Match 0% Maximum Match 100%	RD MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
	Listing first 45 summaries	RX
		RA Xu J., Bjursell M.K., Hämrod J., Deng S., Carmichael L.K.,
		RA Chiang H.C., Hooper L.V., Gordon J.I.; RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
		RL Science 299:2074-2076 (2003).
		DR BEMBL; AE016933; AAO76945.1; -
		DR GO:GO003844; EC: catalytic activity; IEA.
		DR InterPro; IPRE00105; Myb DNA binding.
		DR InterPro; IPRE01375; Peptidase_S9.
		DR InterPro; IPRE06032; Ribosomal_S12_23.
		DR InterPro; IPRE00079; Ser ester.
		DR InterPro; IPRE01042; TolB_C.
		DR Pfam; PF00328; Peptidase_SP_1.
		DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
		DR PROSITE; PS00055; RIBOSOMAL_S12; UNKNOWN_1.
		KW Complete proteome.
		SQ SEQUENCE 840 AA; 94535 MW; 789B533158D6A2; CRC844;
		Query Match 83.3%; Score 50; DB 2; Length 840; Best Local Similarity 66.7%; Prod. No. 31; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
		QY 1 XGKGKGKGW 9
		Db 716 TSIWGEIW 724
		RESULT 2
		ID Q64QS3 PRELIMINARY; PRT; 845 AA.
		AC Q64QS3; DT 25-OCT-2004 (TrEMBLrel. 28, Created) DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
		DE Putative alanyl dipeptidyl peptidase.
		GN ORNames-BP3415;
		OS Bacteroides fragilis.
		OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;



[1] OC Thermoplasmataceae; Thermoplasma.  
 QX NCBI\_TAXID=2303;  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=DSM 1728;  
 RC MEDLINE=20479972; PubMed=11029001; DOI=10.1039/35035069;  
 RX Ruepp A., Graml W., Santos-Martinez M.-L., Kortke K.K., Volker C.,  
 RA Newes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum." ;  
 RT Nature 407:508-513 (2000).  
 RL DR EMBL; AL445063; CAC11406.1;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 748 AA; 82250 MW; 73A030A0E217EE39 CRC64;  
 Query Match Similarity 80.0%; Score 48; DB 2; Length 748;  
 Best Local Similarity 66.7%; Pred. No. 56;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 XGYWGKGYW 9  
 Db 530 SNWKGKAYW 538

RESULT 7  
 Q759B3 PRELIMINARY; PRT; 1293 AA.  
 ID Q759B3  
 AC Q759B3;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last Sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE ADR364WD.  
 GN ORFNames=ADR364W;  
 OS Ashbya gossypii (Yeast)  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetidae; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Eremothecium.  
 NCBI\_TAXID=31169;  
 RN (1).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 10895;  
 RA Voegeli S.E., Brachat S., Dietrich F.S., Lerch A., Gaffney T.,  
 RA Philippo P.;  
 RA Submitted (SBP=2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AB016817; AAS52284.1; -;  
 DR AGD; ADR364W; -;  
 DR InterPro; IPR0005977; DNA ligase.  
 DR InterPro; IPR000886; ER\_Farget\_S.  
 DR PROSITE; PS00697; DNA\_LIGASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
 SQ SEQUENCE 1293 AA; ER\_TARGET; UNKNOWN\_1.  
 DR 145770 MW; 2FB177BDDE93B37 CRC64;

Query Match Similarity 80.0%; Score 48; DB 2; Length 1293;  
 Best Local Similarity 66.7%; Pred. No. 92;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 XGYWGKGYW 9  
 Db 1172 HGTYWEEGYW 1180

RESULT 8  
 KRE5\_YEAST STANDARD; PRT; 1365 AA.  
 ID KRE5\_YEAST  
 AC P22023; Q12190;  
 DT 01-AUG-1991 (Rel. 1.9, Created)  
 DT 01-NOV-1997 (Rel. 3.5, Last sequence update)  
 DT 05-JUL-2004 (Rel. 4.4, Last annotation update)  
 DE Killer toxin-resistance protein 5 precursor.  
 GN Saccharomyces cerevisiae (Baker's Yeast).  
 OS Saccharomyces cerevisiae; OrderedLocusNames=YOR336W;  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetidae; Saccharomycetes;  
 OC Saccharomyces; Saccharomycetaceae; Saccharomyces;  
 OC NCBI\_TAXID=4932;  
 OC NCBI\_TAXID=2303;

[1] OC SEQUENCE FROM N.A.  
 RP RX MEDLINE=90258892; PubMed=2168106;  
 RA Meaden P., Hill K., Wagner J., Slipetz D., Sommer S.S., Bussey H.;  
 RT "The yeast KRE5 gene encodes a probable endoplasmic reticulum protein required for (1->6)-beta-D-glucan synthesis and normal cell growth." ;  
 RL Mol. Cell. Biol. 10:3013-3019 (1990).  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051586; PubMed=8896263;  
 RA Parle-Medermann A.G., Hand N.J., Goulding S.G., Wolfe K.H.;  
 RT "Sequence of 29 kb around the PDR10 locus on the right arm of chromosome XV: similarity to part of Saccharomyces cerevisiae chromosome XV; similarity to part of chromosome I." ;  
 RT Yeast 12:999-1004 (1996).  
 CC -i- FUNCTION: Required for (1->6)-beta-D-glucan synthesis and normal cell growth.  
 CC -i- SUBCELLULAR\_LOCATION: Endoplasmic reticulum lumen.  
 CC -i- SIMILARITY: Some, to D.melanogaster UGGs.  
 CC -i-  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC -i-  
 CC DR M33556; AAA4725.1; -;  
 DR EMBL; Z49821; CAR39981.1; -;  
 DR EMBL; Z75244; CAA39659.1; -;  
 DR PIR; S62066; BVBYKS.  
 DR GermOnline; 143924.  
 DR SGD; S000005863; KRES.  
 DR GO:GO:0003980; F UDP-Glucose:glycoprotein glucosyltransferases. . . ; IDA.  
 DR InterPro; IPR00086; ER TARGET\_S.  
 DR PROSITE; PS00014; ER TARGET; 1.  
 DR Cell wall; Endoplasmic reticulum; Glycoprotein; Signal.  
 DR SIGNAL 1 17 Potential.  
 FT CHAIN 18 1365 Killer toxin-resistance protein 5.  
 FT CARBOHYD 115 115 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 228 228 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 293 293 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 457 457 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 519 519 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 523 523 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 644 644 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 870 870 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 1091 1091 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 1150 1150 N-linked (GlcNAc. . . ) (Potential).  
 FT CARBOHYD 1195 1195 N-linked (GlcNAc. . . ) (Potential).  
 FT SITE 1362 1362 Prevent secretion from ER.  
 FT CONFLICT 582 582 Missing (in Ref. 1).  
 FT CONFLICT 780 794 HLDQNEVPEHFEA -> ILIKMKCQKQNISKAK (in Ref. 1).  
 SQ SEQUENCE 1365 AA; 156476 MW; D0F5851175CC0333 CRC64;  
 SQ Query Match Similarity 80.0%; Score 48; DB 1; Length 1365;  
 SQ Best Local Similarity 66.7%; Pred. No. 96;  
 SQ Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0; PRT; 574 AA.

RESULT 9  
 QY 1 XGYWGKGYW 9  
 DB 1233 DGYWKEGYW 1241  
 AC Q9H5J0; 2B73\_HUMAN STANDARD;  
 DT 10-OCT-2003 (Rel. 4.2, Created)  
 DT 10-OCT-2004 (Rel. 4.2, Last sequence update)  
 DT 05-JUL-2004 (Rel. 4.4, Last annotation update)

DE Zinc finger and BTB domain containing protein 3.	CC Zinc-finger and BTB domain containing protein 3.
OS Homo sapiens (Human).	CC or send an email to license@isb-sib.ch).
OC Burkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	CC
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	CC
[1]	CC
RN SEQUENCE FROM N.A.	CC
RP PubMed=14702039; DOI=10.1038/ng1285;	DR DR Zinc and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,	DR ENBL; AK027045; BAB15636.1; -
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,	DR ENBL; BC025249; AAH2549.1; -
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,	DR Genew; HGNC-22918; ZBTB3.
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,	DR InterPro; IPR000210; BTB_P0Z.
RA Murakami K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,	DR InterPro; IPR007087; Znf_C2H2.
RA Shiratori A., Sudo H., Hosoi T., Karu Y., Kodaira H., Kondo H.,	DR Pfam; PF00651; BTB; 1.
RA Sugawara M., Takahashi M., Kanda K., Yoko T., Furuya T., Kikkawa B.,	DR Pfam; PF00962; zf-C2H2; 2.
RA Ono Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M.,	DR SMART; SM00355; Znf_C2H2; 2.
RA Yamazaki M., Ninomiya K., Ishibashi T., Yanashita H., Murakawa K.,	DR PROSITE; PS00097; BTB; 1.
RA Fujimori K., Tanih H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,	DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
RA Ishitani S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotura T.,	DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
RA Kusano J., Kamehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,	DR DNA-binding; Metal-binding; Nuclear protein; Polymorphism; Repeat; Zinc_finger.
RA Nomura Y., Togoya S., Komai F., Hara K., Takeuchi K., Arita M.,	FT DOMAIN 74 142
RA Imose N., Masaehiro K., Yuki H., Ohshima A., Sasaki N., Aotsuka S.,	FT DOMAIN 227 346
RA Yoshihikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,	FT ZNFING 472 494
RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,	FT ZNFING 500 523
RA Hisamizaki H., Watanabe T., Sugiyama H., Goto Y., Shimizu F., Wakebe H.,	FT VARIANT 574 574
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,	FT FT 1->M (in dbSNP: 544641).
RA Fujimori K., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,	FT FT /FTID:018382.
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,	FT FT /FTID:28C2PF4DBCC4A036 CRC64;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S.,	SQ SEQUENCE 574 AA: 61827 MW; 28C2PF4DBCC4A036 CRC64;
RA Kaitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,	Query Match 76.7%; Score 46; DB 1; Length 574;
RA Matsumura K., Nakakima Y., Mizuno T., Morinaga M., Sasaki M.,	Best Local Similarity 66.7%; Pred. No. 87;
RA Togashi T., Oyama M., Hata H., Watanabe M., Konatsu T.,	Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,	Qy 1 XGYWGKGW 9
RA Okumura K., Nagase T., Nomura N., Kiluchi H., Masuho Y., Yamashita R.,	Db 213 SCHWGRGEW 221
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;	RESULT 10
RT "Complete sequencing and characterization of 21,243 full-length human	Q83VF7 ID PRELIMINARY; PRT; 74 AA.
RT cDNAs."	AC 063VF7 ID PRELIMINARY; PRT; 74 AA.
RT Nat. Genet. 36:40-45 (2004).	AC 063VF7 ID PRELIMINARY; PRT; 74 AA.
RN [2]	AC 063VF7 DT 25-OCT-2004 (TREMBLrel. 28, Created)
RP SEQUENCE FROM N.A.	AC 063VF7 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
RC TISSUE=Skin;	DE Putative lipoprotein.
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242663899;	OS Burkholderia pseudomallei K96243.
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,	OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
RA Altschul S.F., Zeeberg B., Schaefer C.P., Brat N.K.,	OC Burkholderiaceae; Burkholderia.
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	OX NCBI_TaxID:272560;
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	RN [1]
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,	RP SEQUENCE FROM N.A.
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinco P., Frange C.,	RC STRAIN=K96243;
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	RX PubMed=1537794;
RA Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	RA Holden M.T.G., Ritball R.W., Peacock S.J., Cardeno-Tarraga A.M.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,	RA Bentley S.D., Sebbinia M., Thomson N.R., Basen N., Becham I.R.,
RA Fahey J., Helton E., Kettner M., Madan B., Rodriguez S., Sanchez A.,	RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Chevrevach I.,
RA Whiting M., Madan A., Young A.C., Bouffard G.G.,	RA Chillingworth T., Cronin A., Crosson B., Davis P., Deshazer D.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	RA Keith K.B., Maddison P., Moule S., Price C., Quail M.A.,
RA Butterfield Y.S.N., Kiryutkin M.I., Smilus D.E.,	RA Rabbinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.,	RA Songsvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RT "Generation and initial analysis of more than 15,000 full-length human	RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.,
RT and mouse cDNA sequences."	RT "Genomic plasticity of the causative agent of melioidosis,
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	RT Burkholderia pseudomallei;"
CC -1- FUNCTION: May be involved in transcriptional regulation.	RL DR Procs. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).	CC DR EMBL; BX571965; CAH35282.1; -.
CC -1- SIMILARITY: Contains 1 BTB/P0Z domain.	CC KW Lipoprotein.
CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.	CC SQ SEQUENCE 74 AA: 8055 MW; DC8E9521BDEFDA72 CRC64;
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	Query Match 75.8%; Score 45.5; DB 2; Length 74;
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	Best Local Similarity 58.3%; Pred. No. 16;
CC the European Bioinformatics Institute. There are no restrictions on its	Matches 7; Conservative 2; Mismatches 0; Indels 3; Gaps 1;
CC use by non-profit institutions as long as its content is in no way	CC

Qy	1 XGYW--GKCYW 9 :    :   : 60 FGWNSGGGCKYW 71	SEQUENCE FROM N.A. RELMED; AP02819; BAB21080.1; -.	DR EMBL; AP001508; BAB04077.1; -. DR FIR; F03694; F83594. DR GO; GO:0016030; C:membrane; IEA. DR GO; GO:0015635; P:short-chain fatty acid transporter; activity; IEA. DR InterPro; IPR006160; SCFA transp. DR Pfam; PF02667; SCFA transp; 1. KW Complete proteome.
RESULT 11	Q9AXA8 PRELIMINARY; PRT; 198 AA.	SEQUENCE 441 AA; 47239 MW; 5387892FF63302026 CRC64;	SQ SEQUENCE PRELIMINARY; PRT; 1326 AA.
Q9AXA8	AC DT 01-JUN-2001 (TREMBlrel. 17, Created) AC DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update) AC DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)	Query Match Score 45; DB 2; Length 441; Best Local Similarity 75.0%; Pred. No. 96; Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	DR Q6FV20 PRELIMINARY; PRT; 1326 AA. AC Q6FV20 PRELIMINARY; PRT; 1326 AA.
DE	GN Name=Pos01G01.9 OS Eukaryota; Viridiplantae; Streptophytina; Embryophyta; Tracheophyta; OC Spermatophytina; Magnoliophyta; Liliopsida; Poales; Poaceae; OC Ehrhartoideae; Oryzeae; Oryza. OX NCBI_TaxID=39947; [1]	Query Match Score 45; DB 2; Length 441; Best Local Similarity 71.4%; Pred. No. 96; Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	DR 05-JUL-2004 (TREMBlrel. 27, Last sequence update) DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
RN	RP SEQUENCE FROM N.A. RX Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y., RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y., RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H., RA Hosokawa S., Masukawa M., Arikawa K., Chihara Y., Hayashi M., RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C., RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M., RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K., RA Karasawa W., Katajiri S., Kikuta A., Kobayashi N., Kono I., RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y., RA Nakasaki H., Nakashima M., Nakama Y., Nakamura M., RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M., RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K., RA Waki K., Yamagata H., Yamane H., Yosiki S., Yoshihara R., Yukawa K., RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y., RA Yano M., Jiang J., Gojobori T.; RT "The genome sequence and structure of rice chromosome 1."; RL Nature 420:312-316(2002). DR EMBL; AP002819; BAB21080.1; -.	Query Match Score 45; DB 2; Length 198; Best Local Similarity 75.0%; Pred. No. 47; Mismatches 0; Indels 0; Gaps 0;	DR Q6FV20 PRELIMINARY; PRT; 1326 AA; 5387892FF63302026 CRC64;
DR	Gramene; Q9AXA8; -.	Query Match Score 45; DB 2; Length 198; Best Local Similarity 75.0%; Pred. No. 47; Mismatches 0; Indels 0; Gaps 0;	DR Q6FV20 PRELIMINARY; PRT; 1326 AA; 5387892FF63302026 CRC64;
SQ	SEQUENCE 198 AA; 21475 MW; 13F3337BABDB9BEFF0 CRC64;	Query Match Score 45; DB 2; Length 198; Best Local Similarity 75.0%; Pred. No. 47; Mismatches 0; Indels 0; Gaps 0;	DR Q6FV20 PRELIMINARY; PRT; 1326 AA; 5387892FF63302026 CRC64;
Qy	1 XGYWCKGY 8 :    :   : D9 139 LGFWGKGY 146	Query Match Score 45; DB 2; Length 1326; Best Local Similarity 55.6%; Pred. No. 2.6e+02; Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	SQ SEQUENCE 1326 AA; 152292 MW; 694BBE94391ED8DA CRC64;
RESULT 12	Q9KFW3 PRELIMINARY; PRT; 441 AA.	Query Match Score 45; DB 2; Length 1326; Best Local Similarity 55.6%; Pred. No. 2.6e+02; Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	SQ SEQUENCE 1326 AA; 152292 MW; 694BBE94391ED8DA CRC64;
Q9KFW3	ID Q9KFW3; PRELIMINARY; PRT; 441 AA. AC 01-OCT-2000 (TREMBlrel. 15, Created) AC 01-OCT-2000 (TREMBlrel. 15, Last sequence update)	Query Match Score 45; DB 2; Length 1326; Best Local Similarity 55.6%; Pred. No. 2.6e+02; Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	SQ SEQUENCE 1326 AA; 152292 MW; 694BBE94391ED8DA CRC64;
DR	DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)	DR Q9KFW3 PRELIMINARY; PRT; 441 AA.	DR Q9KFW3 PRELIMINARY; PRT; 441 AA.
DE	DE Short-chain fatty acids transporter; GN Name=acots; OrderedLocusNames=BH0358; OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus. OX NCBI_TaxID=86665; [1]	DR Q9KFW3 PRELIMINARY; PRT; 441 AA.	DR Q9KFW3 PRELIMINARY; PRT; 441 AA.
RN	RP SEQUENCE FROM N.A. RC STRAIN=C-125; RX MEDLINE:20512582; PubMed:11058132; DOI=10.1093/nar/28.21.4317; RA Takami H., Nakasone K., Takaki Y., Maeno G., Sesaki R., Masui N., RA Fuji F., Hirama C., Nakamura Y., Ogawara N., Kuhara S., RA Horikoshi K.; RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with <i>Bacillus subtilis</i> ."; RL Nucleic Acids Res. 28:4317-4331(2000).	DR Q9KFW3 PRELIMINARY; PRT; 441 AA. AC 034654; ID YODI_BACSU STANDARD; PRT; 83 AA. DR 10-OCT-2003 (Rel. 42, Created) DT 10-OCT-2003 (Rel. 42, Last sequence update) DT 25-OCT-2004 (Rel. 45, Last annotation update) DE Hypothetical protein yodI.	DR Q9KFW3 PRELIMINARY; PRT; 441 AA. AC 034654; ID YODI_BACSU STANDARD; PRT; 83 AA. DR 10-OCT-2003 (Rel. 42, Created) DT 10-OCT-2003 (Rel. 42, Last sequence update) DT 25-OCT-2004 (Rel. 45, Last annotation update) DE Hypothetical protein yodI.

GN Name=yodI; Synonyms=yola; OrderedLocusNames=BSU19610;  
 OS Bacterium subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TAXID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich S.D.;  
 RT "Sequence analysis of the *Bacillus* subtilis chromosome between  
 the *odhAB* and *sspc* loci cloned in a yeast artificial chromosome.";  
 RT Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 STRAIN=168;  
 RA Ghim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;  
 RT "Sequence analysis of the 30 kb region (182') of the *Bacillus* subtilis  
 chromosome containing the cge cluster.";  
 RT Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
 RA Kunst F., Ogasawara N., Mozer I., Albertini A.M., Alloni G.,  
 RA Azaredo V., Bertuzzo M.G., Bessieres P., Bolotin A., Borchart S.,  
 RA Borriss R., Bourquier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Fouleigre D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Goilley E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holzapfel S., Hosono S., Hull M.F., Itaya M.,  
 RA Jones L.-M., Joris B., Karanam D., Kasahara Y., Klaerr-Bianchard M.,  
 RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S.,  
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
 RA Lazarovic V., Lee S.M., Levine A., Liu H., Masuda S., Mael C.,  
 RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B.,  
 RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
 RA Prescott A.M., Prasecan E., Pujic P., Purnelle B., Rapoport G.,  
 RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
 RA Rose M., Sadie Y., Sato T., Scanlan B., Schleicher S., Schroeter R.,  
 RA Scoffone F., Sekiguchi J., Sekimoto A., Seror S.J., Seror P.,  
 RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
 RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,  
 RA Tognoni A., Tosato P., Uchiyama S., Vandembur M., Vannier F.,  
 RA Vassarotti A., Vari A., Wambutt R., Wedler E., Wedler H.,  
 RA Weitzsaeffer T., Winters P., Wigat A., Yamamoto H., Yamane K.,  
 RA Yoshikawa H., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,  
 PR "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 subtilis.";  
 RL Nature 390:249-256(1997).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EFO15775; AAB72056.1; -.  
 DR EMBL; AF06665; AAB81166.1; -.  
 DR EMBL; 299114; CAB13852.1; -.  
 DR PIR; E69903; E69903.  
 DR BLAST; BG13597; BG13597; yodi.  
 DR InterPro; IPR00891; Transf SH3 like.  
 KW Complete proteome; Hypothetical protein; Transmembrane.  
 FT TRANSMEM 58 80 Potent.A.  
 SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;

Query Match 74.2%; Score 44.5; DB 1; Length 83;

Best Local Similarity 77.8%; Pred. No. 25;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 XGYWGKGYW 9  
 Db 52 YGYWG-GYW 59

Best Local Similarity 77.8%; Pred. No. 25;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 XGYWGKGYW 9  
 Db 108 GGYWG-GYW 115

Search completed: April 19, 2005, 06:42:28  
 Job time : 179 secs

Run on:	April 19, 2005, 05:57:50 ;	Search time 176 Seconds (without alignments)			
Scoring table:	BLOSUM22DX	19.778 Million cell updates/sec			
Perfect score:	60	Gapext 0.5			
Sequence:	xgryngkgwy 9				
Scanned:	2105692 seqs., 386760381 residues				
Total number of hits satisfying chosen parameters:	2105692				
Minimum DB seq length:	0				
Maximum DB seq length:	2000000000				
Post-processing:	Minimum Match 0% Maximum Match 100%				
	Listing first 45 summaries				
Database :	A_Geneseq_16Dec04:*				
	1: geneseqP1980b:*				
	2: geneseqP1990b:*				
	3: geneseqP2000b:*				
	4: geneseqP2001b:*				
	5: geneseqP2002b:*				
	6: geneseqP2003b:*				
	7: geneseqP2004b:*				
	8: geneseqP2004b:*				
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	9	AAV76816	Aav76816 Immunogen
2	59	98.3	8	AAV76814	Aav03714 Fluorine-
3	59	98.3	8	AAV03716	Aav03716 Fluorine-
.4	59	98.3	8	ADG94002	Adg94002 Iodinated
5	59	98.3	8	ADL98013	Peptide h
6	51	85.0	8	AAV76818	Aav76818 Immunogen
7	49	81.7	841	AAB18511	Aab18511 H2 homolo
8	48	80.0	1365	AAR26780	Aar36780 KRES. 7/1
9	48	80.0	1365	AAB72446	Aab72446 Yeast KRE
10	46	76.7	164	AAU16277	Aau16277 Human nov
11	46	76.7	164	ABU53346	Abu53346 Human nov
12	46	76.7	524	ADN93821	Adn93821 Novel hum
13	46	76.7	524	ADN93820	Adn93820 Novel hum
14	45	75.0	8	AAV3715	Aav03715 Fluorine-
15	45	75.0	8	AAV76817	Aav76817 Immunogen
16	45	75.0	8	ADG94005	Adg94005 Immunogen
17	45	75.0	8	ADL98014	Peptide h
18	45	75.0	152	AAB93111	Aab93111 Human pro
19	45	75.0	1090	AAM80087	Aam80087 Human pro
20	45	75.0	1180	ABG19170	Abg19170 Novel hum
21	44.5	74.2	339	ABO78095	Abo78095 Pseudomon
22	44.5	74.2	440	ABU26398	Abu26398 Protein e
23	44	73.3	383	ABO80711	Abos80711 Pseudomon
24	43	71.7	446	ABU23292	Abu23292 Protein e
25	43	71.7	960	ADP18467	Adp18467 S pneumon

conjugates.

PT XX  
PS Claim 21; Page 61; 76pp; English.

CC This sequence represents an immunogenic peptide for a bi-specific antibody. The invention relates to a method of treating or identifying diseased tissues in a patient comprising administering a bi-specific antibody (or fragment) having at least 1 arm (A) that specifically binds a targeted tissue and at least 1 arm (B) that specifically binds a targetable conjugate. The methods and bi-specific antibodies and fusion proteins are useful for pre-targeting methods of diagnosis and therapy. It is advantageous to raise bi-specific antibodies against a targetable conjugate that is capable of carrying at least 1 diagnostic or therapeutic agent. The characteristics of the chelator, metal chelate complex, therapeutic agent or diagnostic agent can be varied to accommodate differing applications without raising new bi-specific antibodies for each new application. The targetable conjugate is selected to elicit sufficient immune responses and also for rapid *in vivo* clearance when used within the bi-specific antibody targeting method.

XX Sequence 9 AA;

Query Match 100.0%; Score 60; DB 3; Length 9;  
Best Local Similarity 100.0%; Prod. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9  
1 ||||| |||  
Db 1 XGYWGKGYW 9

RESULT 2  
AAV03714 standard; peptide; 8 AA.

ID AAV03714;  
XX AC  
DT 08-JUN-1999 (first entry)

DE Fluorine-18 (F-18) labeled peptide 1.  
KW 18F radionuclide; targeting vector; positron emission tomography; F-18;  
KW radiolabeling; thiol; fluorine-18.  
XX Synthetic.

XX Key Location/Qualifiers  
FT Misc-difference 1 /note= "optionally has a free or protected amino acid group"  
FT Misc-difference 2 /note= "D-form residue"  
FT Misc-difference 3 /note= "D-form residue"  
FT Misc-difference 5 /note= "D-form residue; optionally has a free or protected amino acid group"

FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 8 /note= "D-form residue"

XX PN WO9911590-A1.  
PD 11-MAR-1999.  
XX PP 03-SEP-1998; 98WO-US018258.

XX PR 03-SEP-1997; 97US-0057485P.  
PA (IMMU-) IMMUNOMEDICS INC.  
XX Griffiths GL;

XX DR WPI; 1999-228967/19.  
XX PT Radiolabeling thiol-containing peptides with fluorine-18.  
PS Claim 13; Page 15; 22pp; English.

CC The invention relates to a method for incorporating <sup>18</sup>F radionuclide into peptide-containing targeting vectors for use in clinical position emission tomography. Radiolabeling thiol-containing Peptides with <sup>18</sup>F-fluorine-18 (<sup>18</sup>F-18) comprises reacting a peptide comprising a free thiol group with a labeling reagent of formula:  $(\text{CH}_2)_m\text{C}(\text{R}_1)(\text{R}_2)-(\text{CH}_2)_n\text{X}$ , or a fluorinated alkene in which at least one of the two double bonded carbon atoms bears at least one leaving group comprising I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide (optionally substituted by 1-2 alkyl) or 3-sulfonmaleimide; R<sub>1</sub>, R<sub>2</sub> = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH<sub>2</sub>, COOH, OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally substituted by CONH<sub>2</sub>, COOH, OH, sulfonic acid, tertiary amine or quaternary ammonium), COOR', CONR'Cl or COR'; and R' = 1-6C alkyl or phenyl. The method is used for Radio-labeling peptide-containing targeting vectors such as proteins, antibodies, antibody fragments and receptor-targeted peptides for use in routine clinical Positron emission tomography. The method is simple and efficient. The method uses the unique property of the free thiol groups which are rapidly alkylated at neutral pH and moderate temperature. Sequences AAV03714-716 represent examples of F-18 labeled peptides used in the method of detecting a tissue

XX SQ Sequence 8 AA;

Query Match 98.3%; Score 59; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWGKGYW 9  
1 ||||| |||  
Db 1 GYWGKGYW 8

RESULT 3  
AAV03716 standard; peptide; 8 AA.

ID AAV03716;  
XX AC AAV03716;  
XX DT 08-JUN-1999 (first entry)  
XX DE Fluorine-18 (F-18) labeled peptide 3.  
XX KW 18F radionuclide; targeting vector; positron emission tomography; F-18;  
KW radiolabeling; thiol; fluorine-18.  
XX OS Synthetic.

XX Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetylation"  
FT Misc-difference 2 /note= "D-form residue; D-iodo-Tyr"  
FT Misc-difference 3 /note= "D-form residue"  
FT Misc-difference 5 /note= "D-form residue"  
FT Misc-difference 6 /note= "D-form residue; optionally acetylated"  
FT Misc-difference 7 /note= "D-form residue; D-iodo-Tyr"  
FT Misc-difference 8 /note= "D-form residue"  
XX PN WO9911590-A1.  
PD 11-MAR-1999.  
XX PP 03-SEP-1998; 98WO-US018258.  
XX PR 03-SEP-1997; 97US-0057485P.  
PA (IMMU-) IMMUNOMEDICS INC.  
XX Griffiths GL;





FT Misc-difference 8 /note= "D-form iodo-Tyr"  
 FT /note= "D-form residue"  
 XX WO9966951-A2.  
 XX 29-DEC-1999.  
 XX PF 22-JUN-1999; 9W0-US013879.  
 XX PR 22-JUN-1998; 98US-0090142P.  
 XX PR 14-OCT-1998; 98US-0104156P.  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX PI Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;  
 XX DR 2000-160561/14.  
 XX Bi-specific antibodies that bind specific target tissue and targeted  
 PT conjugates.  
 PS Claim 23; Page 61; 76pp; English.  
 CC This sequence represents an immunogenic peptide for a bi-specific  
 CC antibody. The invention relates to a method of treating or identifying  
 CC diseased tissues in a patient comprising administering a bi-specific  
 CC antibody (or fragment) having at least 1 arm (A) that specifically binds a  
 CC targeted tissue and at least 1 arm (B) that specifically binds a  
 CC targetable conjugate. The methods and bi-specific antibodies and fusion  
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.  
 CC It is advantageous to raise bi-specific antibodies against targetable  
 CC conjugate that is capable of carrying at least 1 diagnostic or  
 CC therapeutic agent. The characteristics of the chelator, metal chelate  
 CC complex, therapeutic agent or diagnostic agent can be varied to  
 CC accommodate differing applications without raising new bi-specific  
 CC antibodies for each new application. The targetable conjugate is selected  
 CC to elicit sufficient immune responses and also for rapid in vivo  
 CC clearance when used within the bi-specific antibody targeting method  
 CC Sequence 8 AA;  
 XX SQ Query Match 85.0%; Score 51; DB 3; Length 8;  
 XX Best Local Similarity 87.5%; Pred. No. 1.8e+06;  
 XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX QY 2 GYWKGKGY 9  
 XX 1 GYWKGKGYW 8  
 XX Db AAB18511 standard; protein; 841 AA.  
 XX AC AAB18511;  
 XX DT 15-JAN-2001 (first entry)  
 XX XX H2 homologue of prolyl-tripeptidyl peptidase DPP.  
 XX KW Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;  
 KW gingivitis; periodontitis.  
 XX OS Porphyromonas gingivalis.  
 XX PN WO200052147-A2.  
 XX PD 08-SEP-2000.  
 XX PR 03-MAR-2000; 2000WO-US005551.  
 XX PR 05-MAR-1999; 99US-0123148P.

XX (UGB-) UNIV GEORGIA RES FOUND INC.  
 PA (TRAV/) TRAVIS J.  
 PA (POTB/) POTEMPA J.  
 PA (BANB/) BANBULA A.  
 XX PI Travis J, Potempa J, Banbula A;  
 XX WPI; 2000-594181/56.  
 XX Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful  
 PT for identifying its inhibitor, which is useful for protecting an animal  
 PT from a periodontal disease such as gingivitis and periodontitis.  
 XX Claim 22; Fig 6; 58pp; English.  
 XX PS The present sequence represents a H2 homologue of a prolyl tripeptidyl-  
 CC peptidase (designated DPP) from *Porphyromonas gingivalis*. The prolyl  
 CC tripeptidyl-peptidase has an amidolytic activity, and cleaves a peptide  
 CC bond in a target polypeptide having at least 4 amino acids. This bond is  
 CC between a proline and an amino acid having attached to the alpha-carboxyl group  
 CC end of the proline. The polypeptide is useful for identifying inhibitors  
 CC These inhibitors are then useful for reducing the growth of *Porphyromonas* or  
 CC for protecting an animal from a periodontal disease such as gingivitis  
 CC and periodontitis caused by *Porphyromonas gingivalis*  
 XX Sequence 841 AA;  
 XX SQ Query Match 81.7%; Score 49; DB 3; Length 841;  
 XX Best Local Similarity 66.7%; Pred. No. 1.e+02;  
 XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 XX QY 1 XGTWKGKGY 9  
 XX Db 717 SNTWGSQIW 725  
 XX RESULT 8  
 ID AAB36780 standard; protein; 1365 AA.  
 XX XX AAB36780;  
 AC AC AAB36780;  
 XX DT 16-JUL-1993 (first entry)  
 XX KR55.  
 XX KW Yeast; cell wall; beta-glucan; assembly; pathway; KR1; KR5; growth;  
 KW secretory; O-linked mannose (>6)-beta-glucan; epistasis; morphology;  
 KW hydrophilic; Glycoprotein; COOH-terminal; endoplasmic reticulum; ER;  
 KW reentry signal; antifungal agent.  
 XX OS Saccharomyces cerevisiae.  
 XX PN US5194600-A.  
 XX PD 16-MAR-1993.  
 XX PF 05-MAR-1990; 90US-00488316.  
 XX PR 05-MAR-1990; 90US-00488316.  
 PA (ROYA-) ROYAL INST ADVANCEMENT LEARNING.  
 XX PI Bussey H, Boone C, Sommer SS, Hill K, Meaden P;  
 XX DR WP; 1993-109384/13.  
 XX DR N-PSDB; AAQ38899.  
 XX PT New DNA encoding genes which participate in beta-D-glucan assembly - useful  
 PT for producing mutants for in-vivo screening of antifungal agents and  
 PT providing tools for in-vitro screening.

PS Claim 1; Col 38-44; 24pp; English.  
 XX The sequences given in AAB34785 and AAB36780 represent proteins which  
 CC participate in a yeast cell wall beta-glucan assembly pathway. These  
 CC proteins represent KRE1 and KRE5 respectively, and are essential for  
 CC normal cell growth. KRE1 is a Ser/Thr rich protein that is directed into  
 CC the yeast secretory pathway, where it is highly modified, probably  
 CC through addition of O-linked mannose residues. Gene disruption of the  
 CC KRE1 locus leads to a 40% reduced level of cell wall (1->6)-beta-glucan.  
 CC Mutations at KRE5 also caused defects in cell wall (1->6)-beta-glucan  
 CC production and appears to be epistatic to KRE1. KRE5 is a large  
 CC hydrophilic secretory glycoprotein which contains the COOH-terminal  
 CC endoplasmic reticulum (ER) retention signal (His-Asp-Glu-Leu). Deletion  
 CC of the KRE5 gene results in cells with aberrant morphology and extremely  
 CC compromised growth. KRE1 and KRE5 are useful as tools for the in vitro  
 CC screening of anti-fungal agents which inhibit fungi pathogenic to plants  
 CC and animals. The genes can be used to produce mutants for in vivo  
 CC screening of antifungal agents  
 XX Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 2; Length 1365;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 2; Conservative 1; Indels 0; Gaps 0;  
 SQ Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 4; Length 1365;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 2; Mismatches 1; Indels 0; Gaps 0;  
 SQ Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 4; Length 1365;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 2; Mismatches 1; Indels 0; Gaps 0;  
 SQ Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 4; Length 1365;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 2; Mismatches 1; Indels 0; Gaps 0;

Query Match 9 :||| :|||  
 Db 1233 DGYWKEGYW 1241

Query Match 9 :||| :|||  
 Db 1233 DGYWKEGYW 1241

RESULT 9  
 AAB72446 standard; protein: 1365 AA.  
 ID AAB72446  
 AC AAB72446;  
 DT 08-MAY-2001 (first entry)  
 XX Yeast KRE5.  
 XX UGGT; UDP-glucose:glycoprotein Glucosyltransferase; enzyme; KRE5.  
 XX Saccharomyces cerevisiae.  
 OS WO200112845-A1.  
 PN PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.  
 PD PR 24-FEB-2000; 2000US-0184664P.  
 XX PR 02-MAR-2000; 2000US-0186350P.  
 FF PR 16-MAR-2000; 2000US-0189874P.  
 XX PR 17-MAR-2000; 2000US-0190076P.  
 PR PR 18-APR-2000; 2000US-0198123P.  
 XX PR 19-MAY-2000; 2000US-0205515P.  
 PR PR 07-JUN-2000; 2000US-0209467P.  
 PA PR 28-JUN-2000; 2000US-0214866P.  
 XX PR 30-JUN-2000; 2000US-021515P.  
 PR PR 07-JUL-2000; 2000US-0216641P.  
 DR PR 07-JUL-2000; 2000US-0216880P.  
 XX PR 11-JUL-2000; 2000US-0217487P.  
 PR PR 11-AUG-2000; 2000US-0225213P.  
 CC PR 14-AUG-2000; 2000US-0225214P.  
 CC PR 14-AUG-2000; 2000US-0225266P.  
 CC PR 14-AUG-2000; 2000US-0225267P.  
 CC PR 14-AUG-2000; 2000US-0225288P.  
 CC PR 14-AUG-2000; 2000US-0225290P.  
 CC PR 14-AUG-2000; 2000US-0225477P.  
 CC PR 14-AUG-2000; 2000US-0225737P.  
 CC PR 14-AUG-2000; 2000US-0225738P.

RESULT 10  
 AAU16277 standard; protein: 164 AA.

XX ID AAU16277  
 AC AAU16277;  
 XX AC AAU16277;  
 XX DT 07-NOV-2001 (first entry)  
 XX DE Human novel secreted protein, Seq ID 1230.  
 XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytotoxic;  
 KW cardiotonic; vasotropics; cerebroprotective; nootropic; neuroprotective;  
 KW antibiotic; viricide; fungicide; opthalmological; vulnerary;  
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
 KW cardiovascular disorder; virulence; cerebral ischaemia; angiogenesis;  
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
 KW preservative; antiproliferative.  
 XX OS Homo sapiens.  
 XX PN WO200155322-A2.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US001341.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.  
 XX PR 24-FEB-2000; 2000US-0184664P.  
 XX PR 02-MAR-2000; 2000US-0186350P.  
 XX PR 16-MAR-2000; 2000US-0189874P.  
 XX PR 17-MAR-2000; 2000US-0190076P.  
 XX PR 18-APR-2000; 2000US-0198123P.  
 XX PR 19-MAY-2000; 2000US-0205515P.  
 XX PR 07-JUN-2000; 2000US-0209467P.  
 XX PR 28-JUN-2000; 2000US-0214866P.  
 XX PR 30-JUN-2000; 2000US-021515P.  
 XX PR 07-JUL-2000; 2000US-0216641P.  
 XX PR 07-JUL-2000; 2000US-0216880P.  
 XX PR 11-JUL-2000; 2000US-0217487P.  
 XX PR 11-AUG-2000; 2000US-0225213P.  
 CC PR 14-AUG-2000; 2000US-0225214P.  
 CC PR 14-AUG-2000; 2000US-0225266P.  
 CC PR 14-AUG-2000; 2000US-0225267P.  
 CC PR 14-AUG-2000; 2000US-0225288P.  
 CC PR 14-AUG-2000; 2000US-0225290P.  
 CC PR 14-AUG-2000; 2000US-0225477P.  
 CC PR 14-AUG-2000; 2000US-0225737P.  
 CC PR 14-AUG-2000; 2000US-0225738P.

Disclosure: Fig 9; 95pp; English.

The present invention relates to a method for determining the effect of a

CC test sample on UDP-glucose:glycoprotein Glucosyltransferase (UGGT)

CC activity. The method comprises exposing an acceptor substrate for UGGT to

CC a labelled donor in the presence of the test sample and UGGT. The method

CC is useful for determining UGGT activity. In particular, the method is

CC useful in glucosyltransferase assay and kinetics measurement for

CC determining UGGT activity. UGGT is a soluble enzyme of the endoplasmic

CC reticulum which catalyses the addition of a glucose residue onto

PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-022868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-022709P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 05-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0225509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232339P.  
 PR 14-SEP-2000; 2000US-0232339P.  
 PR 14-SEP-2000; 2000US-0232339P.  
 PR 14-SEP-2000; 2000US-0232339P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233033P.  
 PR 14-SEP-2000; 2000US-0233397P.  
 PR 21-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-023498P.  
 PR 26-SEP-2000; 2000US-0235184P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 02-OCT-2000; 2000US-0236316P.  
 PR 02-OCT-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0235935P.  
 PR 20-OCT-2000; 2000US-0235937P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 08-NOV-2000; 2000US-024478P.  
 PR 08-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-024474P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0244516P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0244532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254070P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI: 2001-488783 / 980pp; English.  
 DR N-PsDB; AAS2264.  
 XX  
 CC The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissue, to regenerate tissues and in chemoraxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence represents a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed



XX	WO2004038003-A2.	PI Williams LT, Chu K, Lee S, Hestir K, Beauroang PA, Behrens D;
PN		PI Halenbeck RF, Kotakota S, Lin H, Linnemann T, Pierce K, Wang Y;
XX		PI Wong JGP, Wu G, Zhang H, Zeng C;
PD	06-MAY-2004.	DR WPI; 2004-365511/34.
XX	24-OCT-2003; 2003MO-US033947.	DR N-PSDB; ADN99037.
PR	25-OCT-2002; 2002US-0421061P.	XX
PR	25-OCT-2002; 2002US-0421080P.	PT New nucleic acid molecules, useful in preparing a composition for
PR	25-OCT-2002; 2002US-0421552P.	PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PR	25-OCT-2002; 2002US-0421614P.	PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PR	30-OCT-2002; 2002US-0422177P.	PT ulcerative colitis.
PR	30-OCT-2002; 2002US-0422178P.	XX
PR	15-NOV-2002; 2002US-0426335P.	Claim 14: SEQ ID NO 14211; 532PP; English.
PR	15-NOV-2002; 2002US-0426338P.	XX
PR	15-NOV-2002; 2002US-0426339P.	The invention relates to a nucleic acid molecule comprising a
PR	15-NOV-2002; 2002US-0426430P.	CC polynucleotide sequence or its complement that encodes a polypeptide. The
PR	15-NOV-2002; 2002US-0426916P.	CC nucleic acid is useful in preparing a composition for treating or
PR	27-NOV-2002; 2002US-0429224P.	CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
PR	27-NOV-2002; 2002US-0429275P.	CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
PR	27-NOV-2002; 2002US-0429320P.	CC heart disease or ulcerative colitis. This sequence corresponds to a
PR	27-NOV-2002; 2002US-0429326P.	CC protein of the invention.
PR	27-NOV-2002; 2002US-0429651P.	XX
PR	04-DEC-2002; 2002US-0429651P.	SQ Sequence 524 AA;
PR	04-DEC-2002; 2002US-0430651P.	Query Match 76.7%; Score 46; DB 8; Length 524;
PR	04-DEC-2002; 2002US-0430657P.	Best Local Similarity 66.7%; Pred. No. 2.3e+02;
PR	04-DEC-2002; 2002US-0430663P.	Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
PR	04-DEC-2002; 2002US-0430668P.	Qy 1 XGYWGKGW 9
PR	04-DEC-2002; 2002US-0430684P.	: :
PR	05-DEC-2002; 2002US-0430693P.	Db 163 SGHWGKGW 171
PR	05-DEC-2002; 2002US-0430965P.	RESULT 13
PR	12-DEC-2002; 2002US-0431459P.	ADN99820
PR	12-DEC-2002; 2002US-0433500P.	ID ADN99820 standard; protein; 524 AA.
PR	13-DEC-2002; 2002US-0433166P.	XX
PR	13-DEC-2002; 2002US-0433188P.	AC ADN99820;
PR	03-JAN-2003; 2003US-04362236P.	XX
PR	17-JAN-2003; 2003US-0437914P.	DT 29-JUL-2004 (first entry)
PR	17-JAN-2003; 2003US-0440821P.	XX
PR	18-APR-2003; 2003US-0453700P.	DB Novel human protein sequence #636.
PR	18-APR-2003; 2003US-0463708P.	XX
PR	18-APR-2003; 2003US-0463716P.	KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;
PR	02-MAY-2003; 2003US-0463732P.	KW antibacterial; viricide; antipsoriatic; cytostatic; gene therapy;
PR	02-MAY-2003; 2003US-0467199P.	KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
PR	02-MAY-2003; 2003US-0467201P.	KW early aging; hormonal imbalance; ischemic heart disease;
PR	02-MAY-2003; 2003US-0467230P.	KW ulcerative colitis.
PR	19-MAY-2003; 2003US-0471306P.	XX
PR	22-MAY-2003; 2003US-0471336P.	Homo sapiens.
PR	09-JUN-2003; 2003US-0472420P.	XX
PR	09-JUN-2003; 2003US-0476609P.	PN WO2004038003-A2.
PR	09-JUN-2003; 2003US-0476621P.	XX
PR	09-JUN-2003; 2003US-0476632P.	PD 05-MAY-2004.
PR	08-JUL-2003; 2003US-0476641P.	XX
PR	08-JUL-2003; 2003US-0485217P.	PF 24-OCT-2003; 2003WO-US033947.
PR	08-JUL-2003; 2003US-0485218P.	XX
PR	08-JUL-2003; 2003US-0485223P.	PR 25-OCT-2002; 2002US-0421061P.
PR	08-JUL-2003; 2003US-0485224P.	PR 25-OCT-2002; 2002US-0421080P.
PR	08-JUL-2003; 2003US-0485325P.	PR 25-OCT-2002; 2002US-0421552P.
PR	08-JUL-2003; 2003US-0485359P.	PR 25-OCT-2002; 2002US-0421614P.
PR	14-JUL-2003; 2003US-0484466P.	PR 30-OCT-2002; 2002US-0422177P.
PR	14-JUL-2003; 2003US-0486480P.	PR 30-OCT-2002; 2002US-0422178P.
PR	15-JUL-2003; 2003US-0486891P.	PR 15-NOV-2002; 2002US-042355P.
PR	15-JUL-2003; 2003US-0489696P.	PR 15-NOV-2002; 2002US-042388P.
PR	08-AUG-2003; 2003US-0493341P.	PR 15-NOV-2002; 2002US-042394P.
PR	08-AUG-2003; 2003US-0493370P.	PR 15-NOV-2002; 2002US-042430P.
PR	08-AUG-2003; 2003US-0493573P.	PR 15-NOV-2002; 2002US-0426916P.
PR	08-AUG-2003; 2003US-0493577P.	PR 27-NOV-2002; 2002US-042224P.
XX		PR 27-NOV-2002; 2002US-042275P.
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	PR 27-NOV-2002; 2002US-042326P.
XX		PR 27-NOV-2002; 2002US-042930P.
PA		PR 27-NOV-2002; 2002US-0429651P.



CC phenyl. The method is used for Radiolabeling peptide-containing targeting  
 CC vectors such as proteins, antibodies, antibody fragments and receptor-  
 CC targeted peptides for use in routine clinical positron emission  
 CC tomography. The method is simple and efficient. The method uses the  
 CC unique property of the free thiol groups which are rapidly alkylated at  
 CC neutral pH and moderate temperature. Sequences AAY03714-716 represent  
 CC examples of F-18 labeled peptides used in the method of detecting a  
 CC tissue  
 XX  
 Sequence 8 AA;  
 Query Match Score 45; DB 2; Length 8;  
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 YWGKCYW 9  
 Db 2 YWGCGYW 8

RESULT 15  
 ID AAY76817 standard; peptide; 8 AA.  
 XX  
 AC AAY76817;  
 XX  
 DT 28-APR-2000 (first entry)  
 DE Immunogenic peptide for bi-specific antibody recognition.  
 XX  
 KW Immunogenic peptide; bi-specific antibody; diagnosis; immune response;  
 KW diseased tissue identification; therapy.  
 XX  
 OS Synthetic.  
 XX  
 FH Location/Qualifiers  
 FT Misc-difference 1  
 FT /note= "acetylated; modified with free amino acid group,  
 FT protected amino acid group, chelating agent or a metal-  
 FT chelate complex"  
 FT Misc-difference 2  
 FT /note= "D-form residue"  
 FT Misc-difference 3  
 FT /note= "D-form residue"  
 FT Misc-difference 5  
 FT /note= "D-form residue; modified with free amino acid  
 FT group, protected amino acid group, chelating agent or a  
 FT metal-chelate complex"  
 FT Misc-difference 7  
 FT /note= "D-form residue"  
 FT Misc-difference 8  
 FT /note= "D-form residue"  
 XX WO9966951-A2.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 22-JUN-1999; 99WO-US013879.  
 XX  
 PR 22-JUN-1998; 98US-009012P.  
 PR 14-OCT-1998; 98US-0104156P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 PI Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;  
 XX  
 DR WPI; 2000-160561/14.  
 XX  
 PT Bi-specific antibodies that bind specific target tissue and targeted  
 PT conjugates.  
 XX  
 PS Claim 22; Page 61; 76pp; English.

CC This sequence represents an immunogenic peptide for a bi-specific  
 CC antibody. The invention relates to a method of treating or identifying  
 CC diseased tissues in a patient comprising administering a bi-specific  
 CC antibody (or fragment) having at least 1 arm (A) that specifically  
 CC binds a targeted tissue and at least 1 arm (B) that specifically binds a  
 CC targetable conjugate. The methods and bi-specific antibodies and fusion  
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.  
 CC It is advantageous to raise bi-specific antibodies against a targetable  
 CC conjugate that is capable of carrying at least 1 diagnostic or  
 CC therapeutic agent. The characteristics of the chelator, metal chelate  
 CC complex, therapeutic agent or diagnostic agent can be varied to  
 CC accommodate differing applications without raising new bi-specific  
 CC antibodies for each new application. The targetable conjugate is selected  
 CC to elicit sufficient immune responses and also for rapid in vivo  
 CC clearance when used within the bi-specific antibody targeting method  
 XX  
 Sequence 8 AA;  
 Query Match Score 45; DB 3; Length 8;  
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWGKCYW 9  
 Db 2 YWGCGYW 8

Search completed: April 19, 2005, 06:39:26  
 Job time : 179 secs

This Page Blank (uspto)

1

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

DBM protein - protein search, using sw model

Run on: April 19, 2005, 06:42:37 ; Search time 133 Seconds  
 (without alignments)

22.491 Million cell updates/sec

Title:	SEQ2	Score:	60
Sequence:	1 xgywgkgrw 9		
Scoring table:	BLOSUM62DX		
Gapext:	Gapext 0.5		
Searched:	1421835 seqs, 332370683 residues		
Total number of hits satisfying chosen parameters:	1421835		
Post-processing: Minimum Match 0% Maximum Match 10% Listing First 45 summaries			
Database :			
Published Applications AA:*			
1: /cgn2_6/ptodata/2/pubpa/us07_pubcomb.pep:*			
2: /cgn2_6/ptodata/2/pubpa/pct_new_pub.pep:*			
3: /cgn2_6/ptodata/2/pubpa/us06_pub.pep:*			
4: /cgn2_6/ptodata/2/pubpa/us06_pubcomb.pep:*			
5: /cgn2_6/ptodata/2/pubpa/us07_new_pub.pep:*			
6: /cgn2_6/ptodata/2/pubpa/pctus_pubcomb.pep:*			
7: /cgn2_6/ptodata/2/pubpa/us08_pub.pep:*			
8: /cgn2_6/ptodata/2/pubpa/us08_pubcomb.pep:*			
9: /cgn2_6/ptodata/2/pubpa/us09_pubcomb.pep:*			
10: /cgn2_6/ptodata/2/pubpa/us09_pubcomb.pep:*			
11: /cgn2_6/ptodata/2/pubpa/us09_pubcomb.pep:*			
12: /cgn2_6/ptodata/2/pubpa/us09_new_pub.pep:*			
13: /cgn2_6/ptodata/2/pubpa/us10_pubcomb.pep:*			
14: /cgn2_6/ptodata/2/pubpa/us10_pubcomb.pep:*			
15: /cgn2_6/ptodata/2/pubpa/us10c_pubcomb.pep:*			
16: /cgn2_6/ptodata/2/pubpa/us10_pubcomb.pep:*			
17: /cgn2_6/ptodata/2/pubpa/us11_new_pub.pep:*			
18: /cgn2_6/ptodata/2/pubpa/us11_new_pub.pep:*			
19: /cgn2_6/ptodata/2/pubpa/us60_new_pub.pep:*			
20: /cgn2_6/ptodata/2/pubpa/us60_pubcomb.pep:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**RESULTS**

Result No.	Score	Query Match Length	Description
1	4.6	76.7	US-09-764-864-1230 Sequence 1230, Ap
2	44.5	74.2	US-10-282-122A-1832 Sequence 4832, A
3	4.3	71.7	US-10-282-122A-1832 Sequence 51216, A
4	4.3	71.7	US-09-769-74A-166 Sequence 166, Ap
5	4.3	71.7	US-10-472-95B-4126 Sequence 48126, Ap
6	4.2	70.0	US-10-437-963-115321 Sequence 115321, Ap
7	4.2	70.0	US-10-767-701-36469 Sequence 36469, A
8	4.2	70.0	US-10-141-312-2 Sequence 2, Appli
9	4.1	68.3	US-10-142-319-136 Sequence 136, App
10	4.1	68.3	US-10-437-963-128022 Sequence 128022, A
11	4.1	68.3	US-10-282-122A-18093 Sequence 48093, A
12	4.1	68.3	US-10-425-114-60598 Sequence 64598, A
13	4.1	68.3	US-10-437-963-131752 Sequence 131752, A

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
1	4.6	76.7	164	US-09-764-864-1230 Sequence 1230, Ap
2	44.5	74.2	440	US-10-282-122A-1832 Sequence 4832, A
3	4.3	71.7	446	US-10-282-122A-1832 Sequence 51216, A
4	4.3	71.7	1329	US-09-769-74A-166 Sequence 166, Ap
5	4.3	71.7	1329	US-10-472-95B-4126 Sequence 48126, Ap
6	4.2	70.0	98	US-10-437-963-115321 Sequence 115321, Ap
7	4.2	70.0	141	US-10-767-701-36469 Sequence 36469, A
8	4.2	70.0	911	US-10-141-312-2 Sequence 2, Appli
9	4.1	68.3	50	US-10-142-319-136 Sequence 136, App
10	4.1	68.3	172	US-10-437-963-128022 Sequence 128022, A
11	4.1	68.3	445	US-10-282-122A-18093 Sequence 48093, A
12	4.1	68.3	453	US-10-425-114-60598 Sequence 64598, A
13	4.1	68.3	469	US-10-437-963-131752 Sequence 131752, A

**ALIGNMENTS**

RESULT 1

; Sequence 1230, Application US/09764864 ; Patent No. US20020132753A1 ; GENERAL INFORMATION: ; APPLICANT: Rosen et al. ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies ; FILE REFERENCE: PTZ23 ; CURRENT APPLICATION NUMBER: US/09/764,864 ; CURRENT FILING DATE: 2001-01-17 ; Prior application data removed - consult PALM or file wrapper ; NUMBER OF SEQ ID NOS: 1792 ; SOFTWARE: PatentIn Ver. 2.0 ; SEQ ID NO 1230 ; LENGTH: 164 ; FEATURES: ; TYPE: PRT ; ORGANISM: HOMO sapiens ; NAME/KEY: SITE ; LOCATION: (2) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids ; NAME/KEY: SITE ; LOCATION: (32) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids ; NAME/KEY: SITE ; LOCATION: (34) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids ; NAME/KEY: SITE ; LOCATION: (84) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids ; NAME/KEY: SITE ; LOCATION: (119) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids ; NAME/KEY: SITE ; LOCATION: (138) ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
 LOCATION: (140)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
 LOCATION: (143)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
 LOCATION: (161)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-764-864-1230

Query Match      Score 46; DB 9; Length 164;  
 Best Local Similarity 66.7%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 XGYWGKGYW 9  
 Db      :|:||| | | |  
 211 PGIGW-GIW 218

RESULT 3  
 US-10-282-122A-51216  
 ; Sequence 51216, Application US/10282122A  
 ; Publication No. US20040009125A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: Wang, Liangsu  
 ;   APPLICANT: Zamudio, Carlos  
 ;   APPLICANT: Malone, Cheryl  
 ;   APPLICANT: Haselbeck, Robert  
 ;   APPLICANT: Ohlsen, Kari  
 ;   APPLICANT: Zyskind, Judith  
 ;   APPLICANT: Wall, Daniel  
 ;   APPLICANT: Trawick, John  
 ;   APPLICANT: Carr, Grant  
 ;   APPLICANT: Yamamoto, Robert  
 ;   APPLICANT: Forsyth, R.  
 ;   APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIORITY APPLICATION NUMBER: 60/191,078  
 ; PRIORITY FILING DATE: 2000-03-21  
 ; PRIORITY APPLICATION NUMBER: 60/206,848  
 ; PRIORITY FILING DATE: 2000-05-23  
 ; PRIORITY APPLICATION NUMBER: 60/207,727  
 ; PRIORITY FILING DATE: 2000-05-26  
 ; PRIORITY APPLICATION NUMBER: 60/230,335  
 ; PRIORITY FILING DATE: 2000-09-06  
 ; PRIORITY APPLICATION NUMBER: 60/230,347  
 ; PRIORITY FILING DATE: 2000-09-09  
 ; PRIORITY APPLICATION NUMBER: 60/242,578  
 ; PRIORITY FILING DATE: 2000-10-23  
 ; PRIORITY APPLICATION NUMBER: 60/253,625  
 ; PRIORITY FILING DATE: 2000-11-27  
 ; PRIORITY APPLICATION NUMBER: 60/257,931  
 ; PRIORITY FILING DATE: 2000-12-22  
 ; PRIORITY APPLICATION NUMBER: 60/267,636  
 ; PRIORITY FILING DATE: 2001-02-09  
 ; PRIORITY APPLICATION NUMBER: 60/269,308  
 ; PRIORITY FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 51216  
 ; LENGTH: 446  
 ; TYPE: PRT  
 ; ORGANISM: Bordetella pertussis  
 ; US-10-282-122A-51216

Query Match      71.7%; Score 43; DB 15; Length 446;  
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      1 XGYWGKGYW 9  
 Db      :|:||| | | |  
 25 PGAWGEGFW 33

RESULT 4  
 US-09-769-744A-166  
 ; Sequence 166, Application US/09769744A  
 ; Publication No. US2003013407A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: Le Page, Richard WF  
 ;   APPLICANT: Wells, Jeremy M  
 ;   APPLICANT: Hanniffy, Sean B

NAME/KEY: SITE  
 LOCATION: (140)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
 LOCATION: (143)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
 LOCATION: (161)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-10-282-122A-48322

Query Match      Score 44.5; DB 15; Length 440;  
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1 XGYWGKGYW 9  
 Db      :|:||| | | |  
 25 PGAWGEGFW 33

Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 48322  
 LENGTH: 440  
 TYPE: PRT  
 ORGANISM: Bacteroides fragilis  
 US-10-282-122A-48322

```

; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122NO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 98163336 .3
; PRIOR FILING DATE: 1998-07-27
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 166
; LENGTH: 1329
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-166

Query Match 71.7%; Score 43; DB 10; Length 1329;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy : ||| | :|
Db 886 ISYWSKGWW 894

RESULT 5
US-10-472-928-4826
Sequence 4826, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P02626WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO: 4826
; LENGTH: 1329
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: beta-N-acetylhexosaminidase (strH)
; OTHER INFORMATION: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Peptide motif: WYY motif
; OTHER INFORMATION: Feature of note: WYY motif
; OTHER INFORMATION: Similar to strain R6 sequence 15902101 (0.E+01)
US-10-472-928-4826

Query Match 71.7%; Score 43; DB 17; Length 1329;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy : ||| | :|
Db 886 ISYWSKGWW 894

RESULT 6
US-10-437-963-115321
Sequence 115321, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; FEATURE:
; OTHER INFORMATION: Human DNA Ligase IV
; FILE REFERENCE: PEI42P1D1
; CURRENT APPLICATION NUMBER: US/10/141,132
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 08/461,562
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: PCT/US94/12922
; PRIOR FILING DATE: 1994-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1

```

```

; SEQ ID NO: 2
; LENGTH: 9:11
; TYPE: PRT
; ORGANISM: human
US-10-141-132-2

Query Match    70.0%; Score 42; DB 13; Length 9:11;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy          1 XGYWGKG 7
Db          468 GGYWGKG 474

RESULT 9
US-10-742-379-136
; Sequence 136, Application US/10742379
; Publication No. US20040181033A1
; GENERAL INFORMATION:
; APPLICANT: Han, HO
; APPLICANT: Min, Hosung
; APPLICANT: Boone, Thomas Charles
; TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN
; FILE REFERENCE: A-828 (US)
; CURRENT APPLICATION NUMBER: US/10/742,379
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/435,923
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 634
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 136
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Myostatin Binding Peptide
US-10-742-379-136

Query Match    68.3%; Score 41; DB 16; Length 50;
Best Local Similarity 55.6%; Pred. No. 66;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy          1 XGYWGKGW 9
Db          35 TGHWAGGYW 43

RESULT 10
US-10-437-963-128022
; Sequence 128022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Barzakov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO: 128022
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure

Query Match    68.3%; Score 41; DB 15; Length 445;
Best Local Similarity 75.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy          1 XGYWGKGY 8
Db          137 AGIWKGKY 144

```

RESULT 12 Sequence 64598 Application US/10425114  
 US-10-425-114-64598  
 Sequence 64598 Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E.  
 APPLICANT: Tabaska, Jack E.  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO: 64598  
 LENGTH: 453  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB4172-021-C8\_FLI.pep  
 US-10-425-114-64598

Query Match 68.3%; Score 41; DB 15; Length 453;  
 Best Local Similarity 55.6%; Pred. No. 4.8e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 SEQ ID NO: 64598

Qy 1 XGYWGRGYW 9  
 Db 33 HGMWGRHYW 41

RESULT 13 Sequence 131752 Application US/10437963  
 US-10-437-963-131752  
 Sequence 131752 Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbausk, Brad  
 APPLICANT: Li, Ping  
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO: 131752  
 LENGTH: 469  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_33789C.1.pep  
 US-10-437-963-131752

Query Match 68.3%; Score 41; DB 16; Length 469;  
 Best Local Similarity 66.7%; Pred. No. 5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWGRGYW 9  
 Db 199 GSYGGKXW 207

RESULT 14 Sequence 11 Application US/10171319  
 US-10-171-319-11  
 Sequence 11 Application US/10171319  
 Publication No. US20030157633A1  
 GENERAL INFORMATION:  
 APPLICANT: Arden Patapoutian  
 APPLICANT: Andrea Peier  
 APPLICANT: Peter McIntyre  
 APPLICANT: Stuart Bevan  
 APPLICANT: Chuanzhang Song  
 APPLICANT: Pampush Ganju  
 TITLE OF INVENTION: VANILLOID RECEPTOR-RELATED NUCLEIC ACIDS  
 FILE REFERENCE: 4-3048A  
 CURRENT APPLICATION NUMBER: US/10/171,319  
 CURRENT FILING DATE: 2002-10-24  
 PRIORITY APPLICATION NUMBER: 60/297,835  
 PRIORITY FILING DATE: 2001-06-13  
 PRIORITY APPLICATION NUMBER: 60/351,238  
 PRIORITY FILING DATE: 2002-01-22  
 PRIORITY APPLICATION NUMBER: 60/352,914  
 PRIORITY FILING DATE: 2002-01-29  
 PRIORITY APPLICATION NUMBER: 60/357,161  
 PRIORITY FILING DATE: 2002-02-12  
 PRIORITY APPLICATION NUMBER: 60/381,086  
 PRIORITY FILING DATE: 2002-05-15  
 PRIORITY APPLICATION NUMBER: 60/381,739  
 PRIORITY FILING DATE: 2002-05-16  
 NUMBER OF SEQ ID NOS: 114  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 11  
 LENGTH: 1268  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-171-319-11

Query Match 68.3%; Score 41; DB 14; Length 1268;  
 Best Local Similarity 55.6%; Pred. No. 1.2e+03;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWGRGYW 9  
 Db 152 GGWGRGLW 160

RESULT 15 Sequence 184877 Application US/10424599  
 US-10-424-599-184877  
 Sequence 184877 Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO: 184877  
 LENGTH: 247  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 NAME KEY: unsure  
 LOCATION: (1)..(247)  
 OTHER INFORMATION: unsure at all xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13795C.1.pep  
 US-10-424-599-184877

Query Match 67.5%; Score 40.5; DB 15; Length 247;  
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Tue Apr 19 06:58:38 2005

seq2.rapp

Page 6

Oy 1 XGYNGKGYN 9  
Db :|||: |||  
108 DGIWGR-IW 115

Search completed: April 19, 2005, 06:55:18  
Job time : 133 secs

Result No.	Score	Query	Match	Length	DB ID	Description
1	48	80.0	1365	3	US-09-376-330-18	Sequence 18, Appl
2	44.5	74.2	339	4	US-09-252-391A-26841	Sequence 2641, A
3	44	73.3	383	4	US-09-252-391A-29457	Sequence 29457, A
4	43	71.7	1290	4	US-09-107-433-4399	Sequence 4399, Ap
5	43	71.7	1303	4	US-09-1583-110-5037	Sequence 5037, Ap
6	42	70.0	844	4	US-09-341-105-2	Sequence 2, Appl
7	42	70.0	911	4	US-08-461-562B-2	Sequence 2, Appl
8	42	70.0	922	4	US-09-902-340-12187	Sequence 12387, A
9	41	68.3	142	2	US-08-860-174A-7	Sequence 7, Appl
10	41	68.3	274	2	US-08-860-174A-12	Sequence 12, Appl
11	41	68.3	282	2	US-08-860-174A-10	Sequence 10, Appl
12	41	68.3	320	4	US-09-248-196A-14837	Sequence 14337, A
13	41	68.3	1447	3	US-09-376-330-17	Sequence 17, Appl
14	40	66.7	139	4	US-09-471-276-837	Sequence 837, App
15	40	66.7	588	4	US-09-540-236-1481	Sequence 2481, Ap
16	40	66.7	624	3	US-08-947-965-78	Sequence 78, Appl
17	40	66.7	655	1	US-08-469-202-27	Sequence 27, Appl
18	40	66.7	655	1	US-08-469-202-28	Sequence 28, Appl
19	40	66.7	655	2	US-08-484-434C-34	Sequence 34, Appl
20	40	66.7	655	2	US-08-484-434C-35	Sequence 35, Appl
21	40	66.7	655	4	US-09-384-361-34	Sequence 34, Appl
22	40	66.7	655	4	US-09-384-361-35	Sequence 35, Appl
23	39	65.0	52	4	US-09-513-399C-4456	Sequence 4456, Ap
24	39	65.0	100	4	US-09-513-399C-5031	Sequence 5031, Ap
25	39	65.0	126	4	US-09-902-540-15783	Sequence 15783, A
26	39	65.0	185	4	US-09-583-110-3898	Sequence 2898, Ap
27	39	65.0	196	4	US-09-107-433-3167	Sequence 3167, Ap

## ALIGNMENTS

RESULT 1  
US-09-376-330-18  
; Sequence 18, Application US/09376330  
; Patent No. 63999121  
; GENERAL INFORMATION:  
; APPLICANT: Tessier, Daniel C.  
; APPLICANT: Dignard, Daniel  
; APPLICANT: Bergeron, John J.M.  
; APPLICANT: Thomas, David Y.  
; TITLE OF INVENTION: Method for screening for UDP-glucose-glycoprotein glucosyltransferase (UGGT) activity and nucleic acid encoding for UGGT  
; TITLE OF INVENTION: UDP-glucose-glycoprotein glucosyltransferase (UGGT) activity  
; FILE REFERENCE: 2139-9-US  
; CURRENT APPLICATION NUMBER: US/09/376,330  
; CURRENT FILING DATE: 1999-08-18  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Fast-SEQ for Windows Version 3.0  
; SEQ ID NO: 18  
; LENGTH: 1,365  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: S. cerevisiae KRES  
US-09-376-330-18

Query Match 80.0%; Score 48; DB 3; Length 1365;  
Best Local Similarity 66.7%; Prod. No. 48;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XGTYKGKGYN 9  
Db 1233 DGTYKGKGYW 1241

RESULT 2  
US-09-262-991A-26841  
; Sequence 26841, Application US/09252991A  
; Patent No. 6551195  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenstein et al.  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 10196-136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142

SEQ ID NO: 26841  
 LENGTH: 339  
 TYPE: PRT  
 ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-26841

Query Match 74.2%; Score 44.5; DB 4; Length 339;  
 Best Local Similarity 77.8%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 XGYWKGKGYW 9  
 Db 260 GGYWG-GYW 267

RESULT 3  
 US-09-252-991A-29457  
 Sequence 29457, Application US/09252991A  
 Patent No. 6551795

GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 29457  
 LENGTH: 383  
 TYPE: PRT  
 ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-29457

Query Match 73.3%; Score 44; DB 4; Length 383;  
 Best Local Similarity 66.7%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKGYW 9  
 Db 222 VGTNGGGYW 230

RESULT 4  
 US-09-107-433-4399  
 Sequence 4399, Application US/09107433  
 Patent No. 6800744

GENERAL INFORMATION:  
 APPLICANT: Lynn A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS  
 NUMBER OF SEQUENCES: 5206  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 024354  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: <Unknown>  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: <Unknown>  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,433  
 FILING DATE: 30-JUN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/ 085131

RESULT 5  
 US-09-583-110-5037  
 Sequence 5037, Application US/09583110  
 Patent No. 6698703

GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al.  
 TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus pneumoniae for Diagnostics and Therapeutics  
 FILE REFERENCE: PATH00-07A  
 CURRENT FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US/09/583,110  
 PRIOR FILING DATE: 1998-05-26  
 PRIOR APPLICATION NUMBER: US 09/107,433  
 PRIOR FILING DATE: 1998-06-30  
 PRIOR APPLICATION NUMBER: US 60/085,131  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: US 60/051,553  
 PRIOR FILING DATE: 1997-07-02  
 NUMBER OF SEQ ID NOS: 5322  
 SEQ ID NO 5037  
 LENGTH: 1303  
 TYPE: PRT  
 ORGANISM: *Streptococcus pneumoniae*  
 US-09-583-110-5037

Query Match 71.7%; Score 43; DB 4; Length 1303;  
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKGYW 9  
 Db 860 ISYWSKGW 868

RESULT 6  
 US-09-341-505-2  
 Sequence 2, Application US/09341505  
 Patent No. 6753158

GENERAL INFORMATION:  
 APPLICANT: Jackson, Stephen P

APPLICANT: Critchlow, Susan E  
 TITLE OF INVENTION: Assays, agents, therapy and diagnosis relating to  
 TITLE OF INVENTION: modulation of cellular DNA repair activity

CURRENT APPLICATION NUMBER: US/09/341,505  
 CURRENT FILING DATE: 1999-07-12  
 EARLIER APPLICATION NUMBER: PCT/GB98/00095  
 EARLIER FILING DATE: 1998-01-13  
 EARLIER APPLICATION NUMBER: GB 9700574.8  
 EARLIER FILING DATE: 1997-01-13  
 EARLIER APPLICATION NUMBER: GB 9713131.2  
 EARLIER FILING DATE: 1997-06-20  
 NUMBER OF SEQ ID NOS: 18  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 2  
 LENGTH: 844  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-341-505-2

Query Match      Best Local Similarity 70.0%; Score 42; DB 4; Length 844;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 XGTYWKGK 7  
 Db      401 GGTYWKGK 407

RESULT 7  
 US-08-461-562B-2  
 Sequence 2, Application US/08461562B  
 Patent No. 6455274

GENERAL INFORMATION:  
 APPLICANT: WEI, YING-FEI  
 APPLICANT: HASELTINE, WILLIAM H  
 TITLE OF INVENTION: HUMAN DNA LIGASE IV  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
 STREET: 9410 KEY WEST AVE  
 CITY: ROCKVILLE  
 STATE: MD  
 COUNTRY: US  
 ZIP: 20850

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,562B  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION NUMBER: PCT/US94/12922  
 APPLICATION NUMBER: US/08/461,562B  
 FILING DATE: 08-NOV-1995  
 ATTORNEY AGENT INFORMATION:  
 NAME: A. ANDERS BROOKES  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PF1442P1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 911 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-461-562B-2

Query Match      70.0%; Score 42; DB 4; Length 911;

Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 XGTYWKGK 7  
 Db      468 GGTYWKGK 474

RESULT 8  
 US-09-902-540-1218-7  
 Sequence 12187, Application US/09902540  
 Patent No. 6833447  
 GENERAL INFORMATION:  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Hankle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Wiegand, Roger C.  
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 FILE REFERENCE: 38-10 (15849)B  
 CURRENT APPLICATION NUMBER: US/09/902,540  
 CURRENT FILING DATE: 2001-07-10  
 PRIORITY APPLICATION NUMBER: 60/217,883  
 PRIORITY FILING DATE: 2000-07-10  
 NUMBER OF SEQ ID NOS: 16825  
 SEQ ID NO 12187  
 LENGTH: 922  
 TYPE: PRT  
 ORGANISM: Myxococcus xanthus  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1). (922)  
 OTHER INFORMATION: unsure at all Xaa locations

US-09-902-540-1218-7

Query Match      Best Local Similarity 70.0%; Score 42; DB 4; Length 922;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 XGTYWKGW 9  
 Db      730 GGTYWGENYW 738

RESULT 9  
 US-08-860-174A-7  
 Sequence 7, Application US/08860174A  
 Patent No. 5989830  
 GENERAL INFORMATION:  
 APPLICANT: DAVIS, Paul James  
 APPLICANT: VAN DER LOOT, Cornelis Paul Erik  
 APPLICANT: VERHOIJEN, Martine Elisa  
 APPLICANT: WILSON, Steve  
 TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT  
 TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP  
 STREET: 1100 New York Avenue, N.W.  
 CITY: WASHINGTON, D.C.  
 STATE:  
 COUNTRY: UNITED STATES  
 ZIP: 20005-3918  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/860,174A  
 FILING DATE: June 16, 1997  
 PRIORITY APPLICATION DATA:  
 PRIORITY NUMBER: EP 95307332.7  
 APPLICATION NUMBER: EP 95307332.7

Query Match    1 XGYWKGYW 9  
 Best Local Similarity    66.7%; Score 41; DB 2; Length 142;  
 Matches    6; Conservative    3; Mismatches    0; Indels    0; Gaps    0;  
 Db    121 HGYXKGKF 129

RESULT 10  
 US-08-860-174A-12  
 Sequence 12, Application US/08860174A  
 Patent No. 5989830  
 GENERAL INFORMATION:  
 APPLICANT: DAVIS, Paul James  
 APPLICANT: VAN DER LOGT, Cornelis Paul Erik  
 APPLICANT: VERHOEIJEN, Martine Elisa  
 APPLICANT: WILSON, Steve  
 TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT  
 MOLECULE TYPE: linear  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP  
 STREET: 1100 New York Avenue, N.W.  
 STREET: 9th Floor, East Tower  
 CITY: WASHINGTON, D.C.  
 STATE:  
 COUNTRY: UNITED STATES  
 ZIP: 20005-3918

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text  
 SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/860,174A  
 FILING DATE: June 16, 1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: EP 95307332.7  
 FILING DATE: October 16, 1995  
 APPLICATION NUMBER: PCT/EP/96/03605  
 FILING DATE: August 14, 1996  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 282 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-860-174A-10

Query Match    1 XGYWKGYW 9  
 Best Local Similarity    66.7%; Score 41; DB 2; Length 282;  
 Matches    6; Conservative    3; Mismatches    0; Indels    0; Gaps    0;

RESULT 12  
 US-09-248-796A-14837  
 Sequence 14837, Application US/09248796A  
 Patent No. 6741137  
 GENERAL INFORMATION:  
 APPLICANT: Keith Weinstock et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 SEQUENCE CHARACTERISTICS:  
 FILE REFERENCE: 107196.132  
 CURRENT APPLICATION NUMBER: US/09/248-796A  
 PRIORITY NUMBER: 14837  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 14837  
 LENGTH: 320  
 TYPE: PCT  
 ORGANISM: Candida albicans  
 US-09-248-796A-14837

Query Match    1 XGYWKGYW 9  
 Best Local Similarity    55.6%; Score 41; DB 4; Length 320;  
 Matches    5; Conservative    4; Mismatches    0; Indels    0; Gaps    0;

RESULT 11  
 US-08-860-174A-10  
 Sequence 10, Application US/08860174A

Qy 1 XGYWKGYW 9  
 Db :|||:  
 Db 287 LGWFKGFW 295

RESULT 13  
 US-09-376-330-17  
 ; Sequence 17, Application US/09376330  
 ; Patent No. 6399321.  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tessier, Daniel C.  
 ; APPLICANT: Dignard, Daniel.  
 ; APPLICANT: Bergeron, John J.-M.  
 ; APPLICANT: Thomas, David Y.  
 ; TITLE OF INVENTION: Method for screening for  
 ; TITLE OF INVENTION: UDP-glucosidase/glycoprotein Glucosyltransferase (UGGT) activity  
 ; FILE REFERENCE: 21139-9"US"  
 ; CURRENT FILING DATE: 1999-08-18  
 ; CURRENT APPLICATION NUMBER: US/09/376,330  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO: 17  
 ; LENGTH: 1447  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: S. pombe UGGT  
 US-09-376-330-17

Query Match Score 41; DB 3; Length 1447;  
 Best Local Similarity 71.4%; Pred. No. 5e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YWGKGYW 9  
 Db :|||:  
 Db 1297 FWKKGYW 1303

RESULT 14  
 US-09-471-276-837  
 ; Sequence 837, Application US/09471276  
 ; Patent No. 6822072.  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Duclert A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6822072  
 ; FILE REFERENCE: GENSET-025CP1  
 ; CURRENT APPLICATION NUMBER: US/09/471,276  
 ; CURRENT FILING DATE: 1999-12-21  
 ; EARLIER APPLICATION NUMBER: 09/057,719  
 ; EARLIER FILING DATE: 1998-04-09  
 ; EARLIER APPLICATION NUMBER: 09/069,047  
 ; EARLIER FILING DATE: 1998-04-28  
 ; EARLIER APPLICATION NUMBER: PCT/IB99/00712  
 ; NUMBER OF SEQ ID NOS: 1622  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO: 837  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -19..-1

US-09-471-276-837

Query Match Score 40; DB 4; Length 139;  
 Best Local Similarity 71.4%; Pred. No. 77;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

This Page Blank (uspto)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 18, 2005, 19:35:54 ; Search time 15.4419 Seconds  
(without alignments)

49.847 Million cell updates/sec

Title: SEQ5

Perfect score: 66

Sequence: 1 cywgccyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: PIR1:\*

2: pIR2:\*

3: pIR3:\*

4: pIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	69.7	554	A28716	saposin precursor
2	45.5	68.9	269	B96704	hypothetical prote
3	45.5	68.9	283	2 A49739	hypothetical prote
4	42	63.6	48	1 KJH048	relaxin - horse (f
5	42	63.6	143	2 I47053	relaxin B, C and A
6	42	63.6	342	2 T31757	hypothetical prote
7	42	63.6	347	2 T31755	hypothetical prote
8	42	63.6	351	2 T31758	hypothetical prote
9	42	63.6	473	2 I54210	N-acetylgalactosam
10	42	63.6	533	1 KJH048	N-acetylgalactosam
11	42	63.6	863	2 D70770	probable glycopro
12	41	62.1	83	2 B72392	hypothetical prote
13	41	62.1	280	2 G84839	late embryogenesis
14	41	62.1	441	2 H96988	integral membrane
15	41	62.1	589	2 A29476	mucaricin acetyl
16	41	62.1	589	2 B29514	mucaricin acetyl
17	41	62.1	590	2 S01114	mucaricin acetyl
18	41	62.1	590	2 S47572	hypothetical prote
19	41	62.1	590	2 AF1810	hypothetical prote
20	41	62.1	741	2 S73827	hypothetical prote
21	41	62.1	2970	2 T08839	polyprotein - marm
22	40	60.6	113	2 PH1663	Ig heavy chain V r
23	40	60.6	130	1 A69894	hypothetical prote
24	40	60.6	130	2 T44809	hypothetical prote
25	40	60.6	396	2 T04561	probable phospho
26	40	60.6	415	2 H86204	short-chain fatty
27	40	60.6	456	2 F83694	n-acetylglucosamin
28	40	60.6	456	2 T40367	hypothetical prote
29	40	60.6	499	2 T36462	nitrogenase molybd

## ALIGNMENTS

## RESULT 1

A28716 saposin precursor - rat

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component of sphingolipid activator protein A2; sulfated glycoprotein 1; sulfate sulfatase

N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C;Species: Rattus norvegicus (Norway rat)

C;Accession: A28716

R;Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.

Biochemistry 27, 4557-4564, 1988

A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat

A;Reference number: A28716; PMID: 3048385

A;Molecule type: mRNA

A;Residues: 1-554 <COL>

A;Cross-references: UNIPROT:P10960; GB:M19936; NID:G206904; PID:AAA42136.1; PID:G206909

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by Edman sequencing.

C;Function:

A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them active.

A;Pathway: sphingolipid catabolism

A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucuronidase

A;Note: saposin B (SAP-1) activates hydrolysis of Galactosylceramide sulfate by arylsphingomyelin phosphodiesterase

C;Superfamily: saposin repeat homology

C;Keywords: alternative splicing; Glycoprotein; lysosomal storage disease; lysozyme; sphingomyelin

F;1-17/554/Product: prosaposin #status Predicted <SIG>

F;17-554/Domain: prosaposin #status Predicted <PRO>

F;55-148/Domain: saposin repeat homology <SAPI>

F;60-143/Domain: saposin A #status Predicted <SAPA>

F;189-280/Domain: saposin repeat homology <SA2>

F;194-273/Product: saposin B #status Predicted <SAB1>

F;306-397/Domain: saposin repeat homology <SA3>

F;310-389/Domain: saposin C #status Predicted <SACP>

F;431-512/Domain: saposin repeat homology <SA4>

F;437-514/Domain: saposin D #status Predicted <SDAD>

F;63-138, 66-132, 94-106, 439-512, 442-506, 470-510/Disulfide bonds: #status Predicted

F;80, 214, 331, 456/Binding site: carbohydrate (Asn) (covalent) #status Predicted

F;197-270, 200-264, 229-240, 314-387, 317-381, 345-356/Disulfide bonds: #status Predicted

Query Match 69.7%; Score 46; DB 1; Length 554;

Best Local Similarity 75.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CYWCGYW 8

Db 525 CYWCGYW 532

RESULT 2

E96704 Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C;Accession: E96704  
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sekano, H.  
 A;Cross-references: UNIPROT:Q9C9X6; GB:AE005173; NID:96553884; PIDN:AAF16550.1; GSPDB:GN  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-269 <STO>  
 A;Cross-references: UNIPROT:Q9C9X6; GB:AE005173; NID:96553884; PIDN:AAF16550.1; GSPDB:GN  
 C;Genetics:  
 A;Map position: 1

Query Match 68.9% Score 45.5; DB 2; Length 269;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CY-WGCGCW 8  
 Db 190 CYRWGCGGW 198

RESULT 3  
 HE4430 hypothetical protein At2g01930 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: H84430  
 R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; ius, D.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10517197

A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-283 <STO>  
 A;Cross-references: UNIPROT:Q9S1K0; GB:AE002093; NID:96598329; PIDN:AAF18588.1; GSPDB:GN  
 C;Genetics:  
 A;Gene: At2g01930  
 A;Map position: 2

Query Match 68.9% Score 45.5; DB 2; Length 283;  
 Best Local Similarity 77.8%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CY-WGCGCW 8  
 Db 204 CYRWGCGGW 212

RESULT 4  
 A49739 relaxin - horse (fragments)  
 C;Species: Equus caballus (domestic horse)  
 C;Date: 07-Apr-1994 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
 C;Accession: B49739; #49739  
 R;Stewart, D.R.; Nevins, B.; Hadas, E.; Vandlen, R.  
 A;Title: Affinity purification and sequence determination of equine relaxin.  
 A;Reference number: A49739; MUID:2055195  
 A;Accession: B49739

A;Molecule type: protein  
 A;Residues: 1-28 <STO>  
 A;Cross-references: UNIPROT:P222969  
 A;Accession: A49739  
 A;Molecule type: protein  
 A;Residues: 29-48 <STO>  
 C;Superfamily: insulin  
 C;Keywords: hormone; pyroglutamic acid  
 F;1-28/48/Product: relaxin #status experimental <MAT>  
 F;1-28/Domain: chain B #status experimental <CHB>  
 F;29-48/Domain: chain A #status experimental <CHA>  
 F;1,29/Modified site: pyrrolidine carboxylic acid (Gln) (in mature form) #status experiment  
 F;10-34/22-39,35-48/Disulfide bonds: #status predicted

Query Match 63.6% Score 42; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYWGC 5  
 Db 35 CYWGC 39

RESULT 5  
 I47053 relaxin B,C and A chains - horse (fragment)  
 C;Species: Equus sp.  
 C;Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 16-Jul-1999  
 C;Accession: I47053  
 R;Klonisch, T.; Ryan, P.L.; Yamashiro, S.; Porter, D.G.  
 Biol. Reprod. 52, 1307-1315, 1995  
 A;Title: Partial complementary deoxyribonucleic acid cloning of equine relaxin messenger  
 A;Reference number: I47053; MUID:95359320; PMID:7543295  
 A;Accession: I47053  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Accession: 1-143  
 C;Superfamily: insulin  
 C;Keywords: relaxin  
 F;1-143/KLO>  
 Query Match 63.6% Score 42; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYWGC 5  
 Db 138 CYWGC 142

RESULT 6  
 T311757 hypothetical protein C07G3.4 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T311757  
 R;Geisel, C.; Wamsley, P.  
 submitted to the EMBL Data Library, July 1997  
 A;Description: The sequence of C. elegans cosmid C07G3.  
 A;Reference number: Z21050  
 A;Accession: T311757  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-342 <GBI>  
 A;Cross-references: UNIPROT:AF016432; EMBL:AAF65382.1; GSPDB:GN00023; CESP:  
 A;Experimental source: strain Bristol N2; clone C07G3  
 C;Genetics:  
 A;Gene: C07G3.4  
 A;Map Position: 5  
 A;Introns: 66/3; 122/2; 167/2; 214/3; 319/1  
 Query Match 63.6% Score 42; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYNGC 5  
Db 97 CYNGC 101

**RESULT 7**  
T31755 hypothetical protein C07G3.6 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
R;Geisel, C.; Wamley, P.  
A;Description: The sequence of *C. elegans* cosmid C07G3.  
A;Accession: T31755  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-347 <GET>  
A;Cross-references: UNIPROT:O16323; EMBL:AF016432; PIDN:AAB65378.1; GSPPDB:GN00023; CESP:  
C;Genetics:  
A;Map Position: 5  
A;Introns: 66/3; 122/2; 167/2; 214/3; 319/1  
Query Match 63.6%; Score 42; DB 2; Length 347;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CYNGC 5  
Db 97 CYNGC 101

**RESULT 8**  
T31758 hypothetical protein C07G3.3 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
R;Geisel, C.; Wamley, P.  
A;Description: The sequence of *C. elegans* cosmid C07G3.  
A;Accession: T31758  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-351 <GET>  
A;Cross-references: UNIPROT:O16326; EMBL:AF016432; PIDN:AAB65377.1; GSPPDB:GN00023; CESP:  
C;Genetics:  
A;Gene: CESP:C07G3.3  
A;Map Position: 5  
A;Introns: 66/3; 122/2; 167/2; 214/3; 319/1  
Query Match 63.6%; Score 42; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CYNGC 5  
Db 97 CYNGC 101

R;Kunieda, T.  
Genomics 29, 582-587, 1995  
A;Title: Mucopolysaccharidoses type VI in rats: Isolation of cDNAs encoding arylsulfatase A.  
A;Reference number: 154210; MUID:96121368; PMID:8573749  
A;Accession: 154210  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-473 <RES>  
A;Cross-references: UNIPROT: P50430; GB:D49434; NID:91065603; PIDN:BA08412.1; PID:91089  
C;Genetics:  
A;Gene: ARSB  
C;Superfamily: animal sulfatase  
C;Keywords: sulfuric ester hydrolase  
F;31/Modified site: 3'-oxoalanine (Cys) #status predicted

Query Match 63.6%; Score 42; DB 2; Length 473;  
Best Local Similarity 85.7%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YNGCGYW 8  
Db 384 YPGGGYW 390

**RESULT 10**  
KJHUBAB  
N-acetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor [validated] - human  
N;Alternate names: arylsulfatase B (ASB); chondroitinase; G4S; N-Carboxy-  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence revision 27-Oct-1995 #text change 09-Jul-2004  
C;Accession: S35990; S33307; A35078; A45659; A42449; B42449; I54217; A56865  
submitted to the EMBL Data Library, March 1993  
A;Reference number: S35990  
A;Accession: S35990  
A;Molecule type: DNA  
A;Residues: 1-533 <PPT>  
A;Cross-references: UNIPROT:P15848; EMBL:X72735; NID:9289009; PIDN:CAA51272.1; PID:9825  
R;Modaresi, S.; Rupp, K.; von Figura, K.; Peters, C.  
BioL. Chem. Hoppe-Seyler 374, 327-335, 1993  
A;Title: Structure of the human arylsulfatase B gene  
A;Reference number: S33307  
A;Accession: S33307  
A;Molecule type: DNA  
A;Residues: 1-104 <MOD>  
A;Cross-references: EMBL:X72735; EMBL:X72736; EMBL:X72737; EMBL:X72738; EMBL:X72739; EM  
A;Note: the enzyme is referred to as EC 3.1.6.9  
R;Peters, C.; Schmidt, B.; Rommerskirch, W.; Rupp, K.; Zuehlsdorf, M.; Vingron, M.; Mey  
J. Biol. Chem. 265, 3374-3381, 1990  
A;Title: Phylogenetic conservation of arylsulfatases. CDNA cloning and expression of hu  
A;Reference number: A35078  
A;Accession: A35078  
A;Molecule type: mRNA  
A;Residues: 1-357; 'V'; 359-533 <PE2>  
A;Cross-references: GB:J05225; NID:9179076; PIDN:AAA51784.1; PID:9179077  
A;Note: parts of this sequence were determined by protein sequencing  
A;Note: the enzyme is referred to as EC 3.1.6.1  
R;Litjens, T.; Morris, C.P.; Gibson, G.J.; Beckmann, K.R.; Hopwood, J.J.  
Biochem. Int. 24, 209-215, 1991  
A;Title: Human N-acetylgalactosamine-4-sulphatase: protein maturation and isolation of  
A;Reference number: A45659; MUID:92028992; PMID:1930244  
A;Accession: A45659  
A;Molecule type: DNA; Protein  
A;Residues: 1-104 <LT>  
A;Cross-references: GB:S57777; NID:9236697; PIDN:AB119988.1; PID:9236698  
A;Note: sequence extracted from NCBI backbone (NCBIN:5777, NCBIPI:57778)  
**RESULT 9**  
I54210 N-acetylgalactosamine-4-sulfatase (EC 3.1.6.1) precursor - rat (fragment)  
N;Alternate names: arylsulfatase (EC 3.1.6.1) B [misidentification]  
C;Species: *Rattus norvegicus* (Norway rat)  
C;Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
C;Accession: I54210

A; Reference number: A42449; MUID:92197625; PMID:1550123  
 A; Accession: A42449  
 A; Molecule type: mRNA  
 A; Residues: 115-116,'R',118 <JIN>  
 A; Cross-references: GB:S90729; NID:9247486; PID:AA021831.1; PID:G247487  
 A; Note: sequence extracted from NCBI backbone (NCBIN:90729, NCBIPI:90731)  
 A; Accession: B42449  
 A; Molecule type: mRNA  
 A; Residues: 234-235,'P',237-238 <JJ2>  
 A; Cross-references: GB:S90736; NID:9247488; PID:AA021832.1; PID:G247489  
 A; Note: sequence extracted from NCBI backbone (NCBIN:90736, NCBIPI:90739)  
 A; Accession: C42449  
 A; Molecule type: mRNA  
 A; Residues: 403-404,'Y',405-407 <JJ3>  
 A; Cross-references: GB:S90743; NID:9247490; PID:AA021833.1; PID:G247491  
 A; Note: the enzyme is referred from NCBI backbone (NCBIN:90743, NCBIPI:90747)  
 A; Note: these mutations give rise to mucopolysaccharidoses type VI, Maroteaux-Lamy disease  
 R; Schmidt, B.; Seimer, T.; Ingendoh, A.; von Figura, K.  
 Cell 82, 271-278, 1995  
 A; Title: A novel amino acid modification in sulfatases that is defective in multiple sulfatases  
 A; Reference number: A51113; MUID:95334208; PMID:7620016  
 A; Contents: annotation; identification of 3-oxoalanine, 2-amino-3-oxopropanoic acid  
 R; Schuchman, E.H.; Jackson, C.E.; Deenick, R.J.  
 Genomics 6, 149-158, 1990  
 A; Title: Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA  
 A; Reference number: I56217; MUID:90152677; PMID:196043  
 A; Accession: I56217  
 A; Status: preliminary; translated from GB/EMBL/DDBJ  
 A; Molecule type: mRNA  
 A; Residues: 1-1357,'V',359-375,'N',377-533 <RES>  
 A; Cross-references: GB:BM32373; NID:9179030; PID:9179030  
 R; Kobayashi, T.; Honke, K.; Jin, T.; Gasa, S.; Miyazaki, T.; Makita, A.  
 Biochim. Biophys. Acta 1159, 243-247, 1992  
 A; Title: Components and proteolytic processing sites of arylsulfatase B from human placenta  
 A; Reference number: A56865; MUID:93003385; PMID:1390929  
 A; Accession: A56865  
 A; Molecule type: protein  
 A; Residues: 41-55;424-425,'X',427-454;466-483 <KOB>  
 A; Experimental source: placenta  
 A; Note: sequence modified after extraction from NCBI backbone  
 A; Note: the fragments shown are the amino ends of the alpha, gamma, and beta chains of the enzyme  
 C; Comment: This enzyme is frequently misidentified as EC 3.1.6.1.  
 C; Genetics:  
 A; Gene: GDB:ARSB  
 A; Cross-references: GB:B112008; OMIM:253200  
 A; Molecule type: protein  
 A; Molecule position: 5q11-5q13  
 A; Intron position: 104/3; 167/1; 230/3; 300/1; 381/2; 405/1; 446/1  
 A; Note: defects in this gene can cause mucopolysaccharidoses type VI, Maroteaux-Lamy disease  
 C; Function:  
 A; Description: hydrolyzes N-acetylgalactosamine-4-sulfate units in chondroitin sulfate  
 C; Superfamily: animal sulfatase  
 C; Keywords: glycoprotein; lysosomal storage disease; sulfatase; sulfatase hydrolase  
 P: 41-40/Domain: signal sequence #status predicted <SIG>  
 P: 41-43/Product: alpha Chain #status predicted <ANAT>  
 P: 466-533/Product: gamma chain #status predicted <BMAT>  
 P: 91/Modified site: 3-oxoalanine (Cys) #status experimental  
 P: 118/279/366,458/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 P: 291/Binding site: carbohydrate (Asn) (covalent) #status absent  
 P: 426/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 63.6%; Score 42; DB 1; Length 533;  
 Best Local Similarity 85.7%; Pred. No. 65; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 1;  
 Qy 2 YWGGYW 8  
 Db 444 YPGGYW 450  
 RESULT 13  
 G84839 late embryogenesis abundant ML7 protein [imported] - Arabidopsis thaliana  
 C; Species: Arabidopsis thaliana (mouse-ear cress)  
 C; Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C; Accession: GB:8839  
 R; Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Talon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, C.; Nature 402, 761-768, 1999  
 A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A; Reference number: A84420; MUID:20083487; PMID:10617197  
 A; Accession: GB:84839

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-280 <STO>  
A:Cross-references: UNIPROT:Q9S7S3; GB:AE002093; NID:q3834196; PIDN:ARC78545.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g11260  
A:Map position: 2

Query Match      62.1%; Score 41; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy      1 CYWGC-GYW 8  
Db      78 CAGGCCGWW 87

**RESULT 14**  
H96968  
integral membrane protein similar to antibiotic resistance protein *B. subtilis* [imported  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: H96968  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, E.V.; Kozin, D.R.  
J.; Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clo  
A:Reference number: A96900; MUTID:21359325; PMID:21359325  
A:Accession: H96968  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-441 <KUR>  
A:Cross-references: UNIPROT:Q97LX0; GB:AE001437; PIDN:AAK78539.1; PID:915023427; GSPDB:G  
C:Genetics:  
A:Gene: CAC0560*

Query Match      62.1%; Score 41; DB 2; Length 441;  
Best Local Similarity 71.4%; Pred. No. 78;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CYWCGY 7  
Db      333 CFWGMGY 339

**RESULT 15**  
A29476  
muscarnic acetylcholine receptor M4 - rat  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004.  
C:Accession: A29476  
R:Braun, T.; Schofield, P.R.; Shivers, B.D.; Pritchett, D.B.; Seburg, P.H.  
Biochem. Biophys. Res. Commun. 149, 125-132, 1987  
A:Title: A novel subtype of muscarinic receptor identified by homology screening.  
A:Reference number: A29476; MUID:8807068; PMID:3120722  
A:Accession: A29476  
A:Molecule type: mRNA  
A:Residues: 1-589 <BR>  
A:Cross-references: UNIPROT:P08483; GB:MI18088; NID:9202657; PID:AAA40659.1; PID:9202658  
A:Experimental source: brain  
C:SuperFamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transme  
P:17-90/Domain: transmembrane #status predicted <TM1>  
P:10-124/Domain: transmembrane #status predicted <TM2>  
P:142-163/Domain: transmembrane #status predicted <TM3>  
P:184-206/Domain: transmembrane #status predicted <TM4>  
P:233-251/Domain: transmembrane #status predicted <TM5>  
P:392-512/Domain: transmembrane #status predicted <TM6>  
P:527-545/Domain: transmembrane #status predicted <TM7>

Query Match      62.1%; Score 41; DB 2; Length 589;  
Best Local Similarity 71.4%; Pred. No. 98;

This Page Blank (uspto)

Scoring table:	BLOSUM62	ALIGMENTS									
Gapop:	10.0 , Gapext: 0.5										
Searched:	1612378 seqs, 512079187 residues										
Total number of hits satisfying chosen parameters:	1612378										
Minimum DB seq length:	0										
Maximum DB seq length:	2000000000										
Post-processing:	Minimum Match 0% Maximum Match 100%										
Database :	UniProt_03: 1: uniprot_sprot: 2: uniprot_trembl: *										
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.										
SUMMARIES											
Result No.	Score	Query Match	Length	DB ID	Description	Match	Best Local Similarity	Score	DB 2;	Length	RT
1	48	72.7	1182	2 Q9ENL3	Q9enl3 colorado ti	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
2	47.5	72.0	134	2 Q7S2D0	Q7s2d0 nurospora	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
3	47	71.2	841	2 Q7MUL8	Q7mul8 porphyromon	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
4	46	69.7	181	2 Q7BWH6	Q7bwh6 diffluovibr	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
5	46	69.7	208	2 Q8KLA42	Q8kla42 rhizobium e	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
6	46	69.7	518	1 SAP_CHICK	013035 gallus galli	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
7	46	69.7	553	2 Q6P744	06p7a4 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
8	46	69.7	554	1 SAP RAT	01960 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
9	45.5	68.9	93	2 Q6Y722	03y722 oryza sativa	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
10	45.5	68.9	269	2 Q9C9X6	09c9x6 arabidopsis	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
11	45.5	68.9	279	2 Q9LDB2	09lde2 arabidopsis	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
12	45.5	68.9	283	2 Q6UL27	Q6ul27 arabidopsis	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
13	45.5	68.9	283	2 Q8SKD0	Q8skd0 arabidopsis	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
14	45.5	68.9	285	2 Q8L9F3	Q8l9f3 arabidopsis	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
15	45	68.2	463	2 Q7MM7	Q7mm7 photobacter	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
16	44	66.7	840	2 Q8A6N9	Q8a6n9 bacteroides	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
17	44	66.7	845	2 Q64QS3	Q64qs3 bacteroides	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
18	43	65.2	78	2 Q6IE10	Q6ie10 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
19	43	65.2	135	2 Q6DM1	Q6dm1 meleagrid h	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
20	43	65.2	145	2 Q6DDM2	Q6ddm2 meleagrid h	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
21	43	65.2	161	2 Q6LQ09	Q6lq09 photobacter	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
22	43	65.2	239	2 Q6P779	Q6p779 oryza sativa	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
23	42	63.6	64	2 Q9UD19	Q9ud19 homo sapiens	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
24	42	63.6	182	1 RELX_HORSE	P22969 equus caballus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
25	42	63.6	185	2 Q6AZ236	Q6az236 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
26	42	63.6	287	2 Q96KT3	Q96kt3 homo sapiens	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
27	42	63.6	314	2 Q8A188	Q8a188 bacteroides	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
28	42	63.6	332	1 CGFL_HUMAN	Q96t75 homo sapiens	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
29	42	63.6	332	1 CGFL_MOUSE	Q8bmj7 mus musculus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
30	42	63.6	332	1 CGFL_RAT	P27587 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
31	42	63.6	342	2 O16325	O16325 caenorhabditis elegans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
32	42	63.6	347	2 O16323	O16323 caenorhabditis elegans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
33	42	63.6	351	2 O16326	O16326 caenorhabditis elegans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
34	42	63.6	473	1 ARSB_RAT	P50430 rattus norvegicus	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
35	42	63.6	533	1 ARSB_HUMAN	P15848 homo sapiens	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
36	42	63.6	863	1 PHSG_MYCTU	Q10639 mycobacterium ulcerans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
37	42	63.6	863	2 Q7U078	Q7u078 mycobacterium ulcerans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
38	42	63.6	871	2 Q7X77	Q7x77 mycobacterium ulcerans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
39	42	63.6	2954	2 Q9698	Q9698 hepatitis g	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
40	42	63.6	2967	2 Q41892	Q41892 hepatitis g	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
41	41	62.1	83	2 Q9NYF1	Q9nyf1 thermotoga maritima	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
42	41	62.1	123	2 Q65XY5	Q65xy5 caenorhabditis elegans	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
43	41	62.1	126	2 Q9VNF1	Q9vnf1 drosophila melanogaster	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
44	41	62.1	225	2 Q817T1	Q817t1 arabiopsis thaliana	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.
45	41	62.1	252	2 P95462	P95462 pectenema	72.7%	100.0%	48;	DB 2;	1182;	SEQUENCE FROM N.A.

RESULT 1	Q9SNL3	ID Q9ENL3 PRELIMINARY; PRT; 1182 AA.
	AC Q9ENL3;	
	DT 01-MAR-2001 (TREMBLrel. 16, Created)	
	DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
	DE VP3.	
	OS Colorado tick fever virus	
	OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.	
	NCBI_TaxID=46839;	
	RN [1]	
	RP SEQUENCE FROM N.A.	
	RC PTRAIN-Florio;	
	RX MEDLINE=20351236; PubMedID=10891182; DOI=10.1006/bbrc.2000.3057;	
	RA Attoui H., Billot F., Bigini P., Cantaloube J.F., de Chesse R., De Micco P., de Lamballerie X.,	
	RT "Sequence determination and analysis of the full-length genome of Colorado tick fever virus, the type species of genus Coltivirus (family Reoviridae)."	
	RT Biochem Biophys Res Commun. 273:1121-1125 (2000).	
	DR AF139159; AAC000618; 1-	
	SQ E1F5F3515B9204C9 CRC64;	
Query	Match	72.7%; Score 48; DB 2; Length 1182;
	Best Local Similarity	100.0%; Pred. No. 72;
	Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTWGCG 6	
	Db 480 CTWGCG 485	

RESULT 2	Q7S2D0	ID Q7S2D0 PRELIMINARY; PRT; 134 AA.
	AC Q7S2D0;	
	DT 01-MAR-2004 (TREMBLrel. 26, Created)	
	DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)	
	DE Predicted protein.	
	GN Name=NCU05348-1;	
	OS Neurospora crassa.	
	OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariiales; Sordariaceae; Neurospora.	
	RC STRAIN-OR7IA;	
	RN [1] NCBI_TaxID=5141;	
	RP SEQUENCE FROM N.A.	
	RA Galagan J.E., Calvo S.E., Borstovich K.A., Selker E.U., Read N.D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B., Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M., Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zeltzer A., Schulte U.,	

RA	Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D., Roy A., Foley K., Nayor M., Thomann N., Barrett R.C., Gneire S., Kamal M., Kamysheva M., Mauceiuk E., Bielke C., Rudd S., Frishman D., Krystofcova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S., Krystofcova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S., Perkins D.J., Pratt C., Catchpole D., Li W., Pratt R.J., Osmani S.A., DeSouza C.C., Macino G., Berglund J., Voelker R., Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R., Natvig D.O., Alex L.A., Mannhaupt G., Freitag R., Paulsen I., Sachse M.S., Landen B.S., Nusbaum C., Birren B., "The Genome Sequence of the Filamentous Fungus Neurospora crassa."; RIL Nature 0:0 (2003).	DR EMBL GenBank DDBJ whole genome shotgun (WGS) entry which is preliminary data.	DR ABX01000427; EAA29556.1; 15465 MW; 39231A0E10C7D57B CRC64;	SEQUENCE 134 AA; ID Q728W6; PRELIMINARY; PRT; 181 AA.	ID Q728W6; PRELIMINARY; PRT; 181 AA.
SQ	Query Match Score 72.0%; Best Local Similarity 53.8%; Matches 7; Qry 1 CY----WGGGY 8; Db 2 CYPEQTWNSCGYW 14	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DT	01-MAR-2004 (TREMBLrel. 26; Last annotation update)	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DE	Conserved domain protein.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
GN	Porphyromonas gingivalis (Bacteroides gingivalis).	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OS	Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales; Porphyromonadaceae; Porphyromonas.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OC		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RC	STRAIN=W83; MEDLINE=22829867; PubMed=12949112; DOI=10.1089/jb.18.5591-5601.2003; Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E., Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J., Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J., Dewhirst F.E., Fraser C.M.; "Complete genome sequence of the oral pathogenic bacterium Porphyromonas gingivalis strain W83.", DR TIGR; AA017178; AAQ66756.1; -.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RN	[1]	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RR	SEQUENCE FROM N. A.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
STRAIN	W83;	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
VERSION	1	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
COMMENT		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RESULT	3	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
ID	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
AC	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DT	01-MAR-2004 (TREMBLrel. 26; Last annotation update)	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DE	Conserved domain protein.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
GN	Porphyromonas gingivalis (Bacteroides gingivalis).	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OS	Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales; Porphyromonadaceae; Porphyromonas.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OC		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RC	STRAIN=W83; MEDLINE=22829867; PubMed=12949112; DOI=10.1089/jb.18.5591-5601.2003; Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E., Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J., Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J., Dewhirst F.E., Fraser C.M.; "Complete genome sequence of the oral pathogenic bacterium Porphyromonas gingivalis strain W83.", DR TIGR; AA017178; AAQ66756.1; -.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RN	[1]	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RR	SEQUENCE FROM N. A.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
STRAIN	W83;	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
VERSION	1	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
COMMENT		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RESULT	3	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
ID	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
AC	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DT	01-MAR-2004 (TREMBLrel. 26; Last annotation update)	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DE	Conserved domain protein.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
GN	Porphyromonas gingivalis (Bacteroides gingivalis).	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OS	Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales; Porphyromonadaceae; Porphyromonas.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OC		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RC	STRAIN=W83; MEDLINE=22829867; PubMed=12949112; DOI=10.1089/jb.18.5591-5601.2003; Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E., Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J., Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J., Dewhirst F.E., Fraser C.M.; "Complete genome sequence of the oral pathogenic bacterium Porphyromonas gingivalis strain W83.", DR TIGR; AA017178; AAQ66756.1; -.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RN	[1]	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RR	SEQUENCE FROM N. A.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
STRAIN	W83;	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
VERSION	1	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
COMMENT		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RESULT	3	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
ID	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
AC	Q7MUI8	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DT	01-OCT-2002 (TREMBLrel. 22; Last sequence update)	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
DE	Conserved domain protein.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
GN	Porphyromonas gingivalis (Bacteroides gingivalis).	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OS	Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales; Porphyromonadaceae; Porphyromonas.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
OC		DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RC	STRAIN=CEN42; MEDLINE=91193195; PubMed=2013564; Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.; "Structural complexity of the symbiotic plasmid of Rhizobium leguminosarum bv. Phaseoli."; J. Bacteriol. 173:2411-2419(1991).	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RN	[1]	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RR	SEQUENCE FROM N. A.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
STRAIN	CEN42;	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
MEDLINE	91193195; PubMed=2013564;	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RT	"Sequence, localization and characteristics of the replicator region of the symbiotic plasmid of Rhizobium etli.";	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RT	Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O., Cevallos M.A., Davila G.; "Sequence, localization and characteristics of the replicator region of the symbiotic plasmid of Rhizobium etli.";	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RT	RESULT 4 Q728W6	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	
RT	71.9 YWGCGY 725	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	DR TIGR; ID Q7MUI8; PRELIMINARY; PRT; 841 AA.	

RN Microbiology 143:2825-2831(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RX MEDLINE=22309397; PubMed=12421308;  
 RA Quintero V.; Cevallos M.A.; Bustos P., Girard L.; Davila G.;  
 RA "A site-specific recombinase (RnQ) is required to exert  
 RTR incompatibility towards the symbiotic plasmid of Rhizobium etli.";  
 RL Mol. Microbiol. 46:103-102 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RA Quintero V.; Bustos P., Girard L.; Rodriguez O., Cevallos M.A.,  
 RA Davila G.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RA Quintero V.; Bustos P., Davila G.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RA Gonzalez V.; Bustos P., Medrano-Soto A.; Ramirez-Romero M.A.,  
 RA Romero D.; Salgado H.; Hernandez-Celis J.C.,  
 RA Quintero V.; Girard L.L.; Rodriguez O., Flores M., Cevallos M.A.,  
 RA Collado-Vides J.; Davila G.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFN42;  
 RA Quintero V.; Bustos P., Davila G.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: U80925; AAC54925; 1;  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 208 AA; 23594 MW; 2DFE1733B000744 CRC64;

<b>Query Match</b> Best Local Similarity 69.7% ; Score 46 ; DB 2 ; Length 208 ; Matches 6 ; Conservative 1 ; Mismatches 1 ; Indels 0 ; Gaps 0 ;	Score 46 ; DB 2 ; Length 208 ; Pred. No. 33 ; 1 CRYGCGW 8 :    : 7 CYFGSGYW 14
---	--

Qy DR

Db DE Proactivator polypeptide Precursor [Contains: Saposin A; Saposin B; Saposin C; Saposin D].

GN Name=PSAP;  
 OS Gallus gallus. (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauvia; Aves; Neognathae; Galliformes; Phasianidae;

NCBI TaxID=9031;  
 RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.

RC TISSUE=Brain, and Liver;  
 RX MEDLINE=981229745; PubMed=9461526;  
 RA Azuma N.; Seo H.-C.; Lie O.; Fu Q.; Gould R.M.; Hiraiwa M.; Burt D.W.;  
 RA Paton J.R.; Morrice D.R.; O'Brien J.S.; Kishimoto Y.;  
 RT "Cloning, expression and map assignment of chicken prosaposin.";  
 RL Biochem. J. 330:321-327(1998).

[2] SEQUENCE FROM N.A.  
 RP Altman N.; Horowitz M.;  
 RA Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place by the sequential action of specific hydrolases. Some of these enzymes require specific low-molecular mass, non-enzymic proteins: the sphingolipids activator proteins (coproteins) (By similarity).  
 CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of Glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and galactosylceramide by beta-galactosylceranidase (EC 3.2.1.46). Saposin C apparently acts by combining with the enzyme and acidic lipid to form an activated complex, rather than by solubilizing the substrate (By similarity).  
 CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galactocerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1 gangliosides by beta-galactosidase (EC 3.2.1.23), and globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22). Saposin B forms a solubilizing complex with the substrates of the sphingolipid hydrolases (By similarity).  
 CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase activator (EC 3.1.4.12) (By similarity).  
 CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).  
 CC -!- PTM: This precursor is proteolytically processed to 4 small peptides, which are similar to each other and are sphingolipid hydrolase activator proteins (By similarity).  
 CC -!- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -!- SIMILARITY: Contains 4 saposin B-type domains.  
 CC --- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license or send an email to license@isb-sib.ch.  
 CC --- EMBL; AB003471; BAA9914.1; -;  
 DR EMBL; AF108656; AAF0589.1; -;  
 DR HSSP; Q92739; 1N69.  
 DR InterPro; IPR003119; SapA.  
 DR InterPro; IPR007856; SapB\_1.  
 DR InterPro; IPR008138; SapB\_2.  
 DR InterPro; IPR008140; SapB\_sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR011001; Saposin like.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF02199; SapA; 2.  
 DR Pfam; PF05184; SapB\_1; 4.  
 DR Pfam; PF03489; SapB\_2; 4.  
 DR PRINTS; PRO1797; SAPOSIN.  
 DR ProdDom; PD000173; SAPB\_sub; 1.  
 DR SMART; SM00162; SAPA; 2.  
 DR SMART; SM00741; SapB; 4.  
 KW Direct protein sequencing; Glycoprotein; GM2-gangliosidosis; Lybosome; Saposin.  
 KW Repeat; Signal; Saposin.  
 FT SIGNAL 1 17 Potential.  
 FT PROPEP 18 60 Saposin D.  
 FT PROPEP 145 193 Saposin A.  
 FT CHAIN 194 276 Saposin B.  
 FT CHAIN 278 305 Saposin C.  
 FT CHAIN 307 387 Saposin D.  
 FT PROPEP 389 398 Saposin.  
 FT PROPEP 399 480 Saposin.  
 FT PROPEP 482 518 Saposin.  
 FT DOMAIN 22 55 Saposin-like type A 1.  
 FT DOMAIN 60 143 Saposin-like type B 1.  
 FT DOMAIN 193 277 Saposin-like type B 2.  
 FT DOMAIN 307 388 Saposin-like type B 3.  
 FT DOMAIN 399 480 Saposin-like type B 4.  
 FT DOMAIN 485 518 Saposin-like type A 2.  
 FT DISULPID 64 139 By similarity.  
 FT DISULPID 67 133 By similarity.  
 FT DISULPID 95 107 By similarity.  
 FT DISULPID 197 273 By similarity.

FT	DISULFID	200	267	By similarity.	DR	InterPro; IPR008138; SapB_2.
FT	DISULFID	229	240	By similarity.	DR	InterPro; IPR008140; SapB_sub.
FT	DISULFID	311	384	By similarity.	DR	InterPro; IPR008373; SapBin.
FT	DISULFID	314	378	By similarity.	DR	InterPro; IPR008139; Saposin.
FT	DISULFID	342	353	By similarity.	DR	InterPro; IPR011001; Saposin_like.
FT	DISULFID	403	476	By similarity.	DR	Pfam; PF02199; SapA_2.
FT	DISULFID	406	470	By similarity.	DR	Pfam; PF03184; SapB_1_3.
FT	DISULFID	434	445	By similarity.	DR	Pfam; PF03489; SapB_2_4.
FT	CARBONYD	81	81	N-linked (GlcNAc . . .) (Potential).	DR	PRINTS; PR0177; SapOSTN.
FT	CARBONYD	214	214	N-linked (GlcNAc . . .) (Potential).	DR	ProDom; PD00132; SapB_sub_1.
FT	CARBONYD	328	328	N-linked (GlcNAc . . .) (Potential).	DR	SMART; SM00162; SAPA_2.
FT	CARBONYD	420	420	N-linked (GlcNAc . . .) (Potential).	SQ	SEQUENCE 553 AA; 61019 MW; 63F3DDSE0C523393 CRC64.
FT	CONFLICT	94	94	R -> T (in Ref. 2).		
FT	CONFICT	486	486	E -> D (in Ref. 2).		
SQ	SEQUENCE	518 AA;	57601 MW;	B803000E91C3963 CRC64;	Query Match	69.7%; Score 46; DB 2; Length 553;
Qy	1 CYWGCGYW	8		Best Local Similarity 69.7%; Pred. No. 75; Mismatches 0; Gaps 0;	PRT	Score 46; Pred. No. 75; Mismatches 0; Gaps 0;
Db	489 CWGPGTGW	496		DB 524 CWGPGYW 531		
<b>RESULT 8</b>						
SAP_RAT						
ID	SAP_RAT			STANDARD:	PRT;	554 AA.
AC	P1050;	Q62841;	Q61190;			
DT	01-JUL-1989 (Rel. 11, Created)					
DT	01-JUL-1989 (Rel. 11, Last sequence update)					
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)					
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)					
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)					
DE	prosaposin.					
GN	Name=Bsap;					
OS	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: OC					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.					
OX	NCBI_TaxID=10116;					
RN	SEQUENCE FROM N.A.					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Prostate;					
RX	Medline=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;					
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., "Sequence update,"					
RA	Klausner R.D., Collier P.S., Wagner L.L., Schaefer C.F., Bhat N.K.,					
RA	Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,					
RA	Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J.J., Hsien F.,					
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,					
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,					
RA	Bosak S.A., Loqueland N.A., Peters G.J., Abramson R.D., Mulahay S.J.,					
RA	Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Fahey J.J., Helton E., Ketteman M., Madan A., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Whiting M., Madan A., Young A.C., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Bouffard G.G., Rodriguez A.C., Grimwood J.W., Schmutz J., Myers R.M., Butterfield Y.S.,					
RA	Krzywinski M.I., Skalska U., Smailus D.B., Schnurch A., Schein J.E., Jones S.J., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences," Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).					
RN	SEQUENCE FROM N.A.					
RC	TISSUE=Prostate;					
RA	Strauberg R.; Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.					
RL	EMBL; BC061759; AY61759.1;					
DR	GO; GO:0005764; C:lysosome; IEA.					
DR	GO; GO:0006629; P:lipid metabolism; IEA.					
DR	GO; GO:0006635; P:spingolipid metabolism; IEA.					
DR	InterPro; IPR003119; SapA_1.					
DR	InterPro; IPR007856; SapB_1.					
DR	EMBL; M142136.1; -.					

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR

EMBL; S81353; AAB36042; -;	RP SEQUENCE FROM N.A.
DR EMBL; S81373; AAB36233; 2;	RA Seasaki T., Matsunoto T., Katayose Y.; Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR PIR; A2816; A28716;	RL
DR HSSP; Q92739; 1N69;	DR AP005644; BAD17478.1; -
DR RGD; 3423; PspB.	KW Hypothetical protein.
DR InterPro; IPR003119; SapA.	SQ SEQUENCE 93 AA; 9536 MW; A206432A3764546F CRC64;
DR InterPro; IPR007856; SapB_1.	
DR InterPro; IPR008138; SapB_2.	
DR InterPro; IPR008144; SapB_sub.	
DR InterPro; IPR008137; Saposin.	Query Match 68.9%; Score 45.5%; Pred. No. 20;
DR InterPro; IPR001001; Saposin like.	Best Local Similarity 63.6%; Mismatches 1; Indels 3; Gaps 1;
DR InterPro; IPR008139; SaposinB.	
PFam; PF02199; SapA; 2.	
PFam; PF05184; SapA; 1.	
PFam; PF03489; SapB; 4.	
DR PRINTS; PR01737; SAPOSIN.	
ProDom; PD001732; SAPOSIN.	
Direct Protein sequencing; Glycoprotein; Repeat; Signal.	
FT SIGNAL 1	RESULT 10
FT CHAIN 16	Q9C9X6 PRELIMINARY; PRT; 269 AA.
FT DOMAIN 17	ID Q9C9X6
FT DOMAIN 554	AC Q9C9X6;
FT DOMAIN 21	DT 01-JUN-2001 (TREMBLrel. 17, Created)
FT DOMAIN 54	DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
FT DOMAIN 61	DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
FT DOMAIN 193	DB Hypothetical protein T23K23.3 (Basic pentacyctine 3).
FT DOMAIN 274	GN Name=T23K23.3; Synonyms=BPC3; ORFNAMES=At1g68120;
FT DOMAIN 310	OS Arabidopsis thaliana (Mouse-ear cress).
FT DOMAIN 391	OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
FT DOMAIN 435	OC Magnoliophyta; eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.
FT DOMAIN 516	OC
FT DOMAIN 521	OC
FT DISULFID 554	OC
FT DISULFID 63	OC
FT DISULFID 138	OC
FT DISULFID 66	OX NCBI_TAXID=1702;
FT DISULFID 132	RN
FT DISULFID 94	RP SEQUENCE FROM N.A.
FT DISULFID 106	RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.J., Wu D., Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.B., Bowman C.L., White O., Niemann W.C., Fraser C.M., SubMITTED (OCT-1999) to the EMBL/GenBank/DBJ databases.
FT DISULFID 197	[1]
FT DISULFID 270	RN
FT DISULFID 200	RA
FT DISULFID 264	RA
FT DISULFID 229	RA
FT DISULFID 240	RA
FT DISULFID 314	RA
FT DISULFID 387	RA
FT DISULFID 317	RA
FT DISULFID 381	RA
FT DISULFID 345	RA
FT DISULFID 356	RA
FT DISULFID 439	RA
FT DISULFID 512	RA
FT DISULFID 442	RA
FT DISULFID 506	RA
FT DISULFID 470	RA
FT CARBOHYD 481	RA
FT CARBOHYD 80	RA
FT CARBOHYD 214	RA
FT CARBOHYD 331	RA
FT CARBOHYD 456	RA
FT CONFLICT 115	RA
FT CONFLICT 299	RA
FT CONFLICT 462	RA
FT CONFLICT 527	RA
FT CONFLICT 536	RA
SEQUENCE 554 AA; 61123 MW;	RA
Query Match 69.7%; Score 46; DB 1; Length 554;	RP SEQUENCE FROM N.A.
Best Local Similarity 75.0%; Pred. No. 75;	RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	RL [3]
QY 1 CYWGGCYW 8	RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
Db 525 CYWGGCYW 532	RN
RESULT 9	RESULT 11
Q6YY22 ID Q9LDE2 PRELIMINARY; PRT; 279 AA.	Q9LDE2 PRELIMINARY; PRT; 279 AA.
AC Q6YY22; DT 05-JUL-2004 (TREMBLrel. 27, Created)	ID Q9LDE2
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)	AC Q9LDE2
DR 05-JUL-2004 (TREMBLrel. 27, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)
DB Hypothetical protein OSJNBB056122.26.	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
OS Oriza sativa (japonica cultivar group).	DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	DE F10B6.5 (TSE21.17) (At1g14680/F10B6_22) (Basic pentacycteine 2).
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
Ehharoideae; Orzyeae; Orzya.	
NCBI_TAXID=39947;	
RN	

Page 6

GN Name=At2g01930;	RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
OS Arabidopsis thaliana (Mouse-ear cress);	RT "Full-length messenger RNA sequences greatly improve genome annotation." ;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	RT annotation." ;
CC euroids II; Brassicales; Rosids;	RT Genome Biol.; 3:RESEARCH0029-RESEARCH0029 (2002).
CX NCBI_TaxID=3702;	RL [2];
RN [1]	RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,	RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Goldmire A.D., Lee J.M., Quach Y., Toriumi M., Wu G., Bowser L.,	RA Feldmann K.,
RA Carrinci P., Chen H., Cheuk Y., Ishizaki Y., Ishida J., Jones T.,	RA Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,	DR EMBL; AY088444; AAC66000.1; -.
RA Miranda M., Narubaka M., Nguyen M., Palm C.J., Sakurai T.,	DR InterPro; IPR010409; DUF1004.1.
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W.,	DR Pfam; PF06217; DUF1004; 1.
RA Theologis A.,	KW Hypothetical protein.
RA Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.	SQ SEQUENCE 285 AA; 31887 MW; B97DB307A81CC504 CRC64;
RN [2]	Query Match 68.9%; Score 45.5%; DB 2; Length 285;
RP SEQUENCE FROM N.A.	Best Local Similarity 77.8%; Pred. No. 51;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,	Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,	Qy 1 CY-WGCGYW 8
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,	Db 206 CYRNGCGW 214
RA Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	RESULT 15
RN [3]	Q7NM7M PRELIMINARY; PRT; 463 AA.
RP SEQUENCE FROM N.A.	ID Q7NM7M
RA Town C.D., Kaul S.,	AC Q7NM7M;
RA Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.	DT 01-MAR-2004 (TREMBLrel. 26, Created)
RN [4]	DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
RP SEQUENCE FROM N.A.	RA Taourit S., Bocs S., Bourreux-Bude C., Chandler M., Charles J.-F.,
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,	DR Dassa B., Deroue R., Deruelle S., Gaudriault S.,
RA Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,	RA Medigue C., Lanols A., Powelli K., Vincent R., Wrigate V.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,	RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.,
RA Carrinci P., Chen H., Cheuk Y., Ishida J., Jones T.,	RA NCBI_TaxID=141679;
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,	DR SEQUENCE FROM N.A.
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,	RC STRAIN="T0101";
RA Sakurai T., Satoh M., Seki M., Shinn P., Southwick A., Shinozaki K.,	DR MEDLINE=22957627; PubMed=14528314;
RA RA Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.	RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
DR EMBL; AY065225; AAL38701.1; -.	RA Taourit S., Bocs S., Bourreux-Bude C., Chandler M., Charles J.-F.,
DR EMBL; AY096550; AAM20200.1; -.	DR Dassa B., Deroue R., Deruelle S., Gaudriault S.,
DR PIR; HB4430; HB4430.	RA Medigue C., Lanols A., Powelli K., Vincent R., Wrigate V.,
DR InterPro; IPR010409; DUF1004.	RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.,
DR Pfam; PF06217; DUF1004; 1.	RA NCBI_TaxID=141679; AA/rel_permeabil.
DR Hypothetical protein.	DR InterPro; IPR002293; AA/rel_permeabil.
SQ SEQUENCE 283 AA; 31648 MW; BAFBFCD92262E07 CRC64;	DR InterPro; IPR004840; AAC_permeabil.
Query Match 68.9%; Score 45.5%; DB 2; Length 283;	DR GO; GO-001650; -.
Best Local Similarity 77.8%; Pred. No. 50;	DR GO; GO-00165021; C:integral to membrane; IEA.
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	DR GO; GO-0005219; F:amino acid-polyamine transporter activity; IEA.
Qy 1 CY-WGCGYW 8	DR GO; GO-0006865; F:amino acid transport; IEA.
Db 204 CYRNGCGW 214	DR GO; GO-0006810; P:transport; IEA.
RESULT 14	DR InterPro; IPR002293; AA/rel_permeabil.
QB1_9F3 ID P0BL9F3 PRELIMINARY; PRT; 285 AA.	DR InterPro; IPR004841; Permease_region.
AC QBL9F3; PROSITE; PS00216; AMINO_ACID_PERMEASE_1.	DR Pfam; PF00324; AA_permeabil.
DT 01-OCT-2002 (TREMBLrel. 22, Created)	DR PROSITE; PS00216; AMINO_ACID_PERMEASE_1.
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)	KW Complete proteome; Transportmembrane; Transport
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)	SQ SEQUENCE 463 AA; 50838 MW; 317CD0ADCP4FCP25 CRC64;
DB3 Hypothetical protein.	Query Match 68.2%; Score 45%; DB 2; Length 463;
OS Arabidopsis thaliana (Mouse-ear cress);	Best Local Similarity 63.6%; Pred. No. 90;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	Matches 7; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
CC euroids II; Brassicales; Rosids;	Qy 2 YWGC---GIW 8
CC NCBI_TaxID=3702;	Db 92 YWGCCLTAWGYW 102
RN [1]	RP SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093376;	Search completed: April 18, 2005, 19:58:36
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,	

Tue Apr 19 06:58:41 2005

seq5.rup

Page 8

Job time : 73.2558 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:18:17 ; Search time 82.0465 Seconds  
 (without alignments)  
 37.711 Million cell updates/sec

Title: SE05  
 Perfect score: 66  
 Sequence: 1 cywgcgyw 8

Scoring table: BL0SUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:  
 1: geneseqp1980s:  
 2: geneseqp1990s:  
 3: geneseqp2000s:  
 4: geneseqp2001s:  
 5: geneseqp2002s:  
 6: geneseqp2003s:  
 7: geneseqp2004s:  
 8: geneseqp2004s:

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:  
 1: geneseqp1980s:  
 2: geneseqp1990s:  
 3: geneseqp2000s:  
 4: geneseqp2001s:  
 5: geneseqp2002s:  
 6: geneseqp2003s:  
 7: geneseqp2004s:  
 8: geneseqp2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	66	100.0	8	2	AAY03715	Aay03715 Fluorine-
2	66	100.0	8	3	AAY66817	Aay76817 Immunogen
3	66	100.0	8	7	ADG94005	Adg94005 Immunogen
4	66	100.0	8	14	ADL98014	Peptide h
5	48	72.7	54	3	AAY66023	Ady66023 Raclomers
6	48	72.7	249	7	ADC33110	Adc33110 Human nov
7	47	71.2	841	3	AAB18511	Aab18511 H2 homolo
8	46	69.7	14	8	ADQ11552	Adq11552 Myostatin
9	46	69.7	50	8	ADQ11655	Adq11655 Myostatin
10	46	69.7	265	6	ABU3201	Abu3201 Protein e
11	45.5	68.9	233	3	AAG40353	Aag40353 Arabidops
12	45.5	68.9	238	3	AAG60115	Aag60115 Arabidops
13	45.5	68.9	252	3	AAG05644	Aag05644 Arabidops
14	45.5	68.9	252	3	AAG0352	Aag0352 Arabidops
15	45.5	68.9	252	3	AAG38268	Aag38268 Arabidops
16	45.5	68.9	257	3	AAG60114	Aag60114 Arabidops
17	45.5	68.9	279	3	AAG38267	Aag38267 Arabidops
18	45.5	68.9	279	3	AAG05643	Aag05643 Arabidops
19	45.5	68.9	279	3	AAG0351	Aag0351 Arabidops
20	45.5	68.9	282	3	AAG38266	Aag38266 Arabidops
21	45.5	68.9	285	3	AAG60113	Aag60113 Arabidops
22	45.5	68.9	286	3	AAG05642	Aag05642 Arabidops
23	45	68.2	8	2	AAY03714	Aay03714 Fluorine-
24	45	68.2	8	2	AAY03716	Aay03716 Fluorine-
25	45	68.2	8	7	ADG94002	Adg94002 Iodinated

RESULT 1  
 ID AAY03715  
 XX standard; peptide; 8 AA.  
 AC AAY03715;  
 XX DT 08-JUN-1999 (first entry)  
 XX DE Fluorine-18 (F-18) labeled peptide 2.  
 XX KW 18F radionuclide; targeting vector; positron emission tomography; F-18;  
 KW radiolabeling; thiol; fluorine-18.  
 XX Synthetic.  
 XX OS Synthetic.  
 XX FH Key  
 FT Misc-difference 1 Location/Qualifiers  
 FT /note= "N-terminal acetylation; optionally has a free or  
 protected thiol group"  
 FT Misc-difference 2  
 FT Misc-difference 3 /note= "D-form residue"  
 FT Misc-difference 4 /note= "D-form residue"  
 FT Misc-difference 5 /note= "D-form residue; optionally has a free or  
 protected thiol group"  
 FT Misc-difference 6 /note= "D-form residue"  
 FT Misc-difference 7 /note= "D-form residue"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 XX PN W09111590-A1.

Claim 14; Page 15; 22pp; English.

.

The invention relates to a method for incorporating  $^{18}\text{F}$  radionuclide into peptide-containing targeting vectors for use in clinical positron emission tomography. Radiolabeling thiol-containing peptides with fluorine-18 ( $\text{F}-18$ ) comprises reacting a peptide comprising a free thiol group with a labeling reagent of formula:  $^{18}\text{F}(\text{CH}_2)_m\text{-CR}_1\text{R}_2\text{-}(\text{CH}_2)_n\text{-X}$ , or a fluorinated alkene in which at least one of the two double bonded carbon atoms bears at least one leaving group comprising I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide (optionally substituted by 1-2 alkyl) or 3-sulfonemaleimide; R<sub>1</sub>, R<sub>2</sub> = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH<sub>2</sub>, COOH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally substituted by CONH<sub>2</sub>), COOH, OH, sulfonic acid, tertiary amine or quaternary ammonium, COOR', CONR'2 or COR'; and R' = 1-6C alkyl or phenyl. The method is used for radiolabeling peptide-containing targeting vectors such as proteins, antibodies, antibody fragments and receptor-targeted peptides for use in routine clinical positron emission tomography. The method is simple and efficient. The method uses the unique property of the free thiol groups which are rapidly alkylated at neutral pH and moderate temperature. Sequences AAY3714-716 represent examples of  $\text{F}-18$  labeled peptides used in the method of detecting a tissue

Sequence 8 AA;

Query Match 100.0%; Score 66; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYWGGYW 8

Db 1 CYWGGYW 8

RESULT 2  
ID AAY76817  
XX AAY76817 standard; peptide: 8 AA.  
AC AC  
XX  
DT 28-APR-2000 (first entry)  
DE Immunogenic peptide for bi-specific antibody recognition.  
XX Immuno-  
KW genic peptide; bi-specific antibody; diagnosis; immune response;  
KW diseased tissue identification; therapy.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1  
FT /note= "acetylated; modified with free amino acid group,  
FT protected amino acid group, chelating agent or a metal-  
FT chelate complex"  
FT Misc-difference 2  
FT /note= "D-form residue"  
FT Misc-difference 3  
FT /note= "D-form residue"  
FT Misc-difference 5  
FT /note= "D-form residue; modified with free amino acid  
FT group, protected amino acid group, chelating agent or a metal-  
FT chelate complex"  
FT Misc-difference 7  
FT /note= "D-form residue"  
FT Misc-difference 8  
FT /note= "D-form residue"  
XX W0966651-A2.  
XX  
PN 29-DEC-1999.  
XX  
PF 22-JUN-1999;  
XX  
PP 99WO-US013879.

XX PR 22-JUN-1998; 98US-0090142P.  
CC PR 14-OCT-1998; 98US-0104156P.  
CC XX  
CC (IMMU-) IMMUNOMEDICS INC.  
CC XX  
CC Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;  
CC XX  
CC WPI; 2000-16056/14.  
CC DR  
CC Bi-specific antibodies that bind specific target tissue and targeted conjugates.  
CC XX  
CC Claim 22; Page 61; 76pp; English.  
CC PS  
CC This sequence represents an immunogenic peptide for a bi-specific antibody. The invention relates to a method of treating or identifying CC diseased tissues in a patient comprising administering a bi-specific CC antibody (or fragment) having at least 1 arm (A) that specifically binds a target tissue and at least 1 arm (B) that specifically binds a targetable conjugate. The methods and bi-specific antibodies and fusion CC proteins are useful for pre-targeting methods of diagnosis and therapy. It is advantageous to raise bi-specific antibodies against a targetable CC conjugate that is capable of carrying at least 1 diagnostic or CC therapeutic agent. The characteristics of the chelator, metal chelate CC complex, therapeutic agent or diagnostic agent can be varied to accommodate differing applications without raising new bi-specific CC antibodies for each new application. The targetable conjugate is selected CC to elicit sufficient immune responses and also for rapid in vivo CC clearance when used within the bi-specific antibody targeting method  
CC XX  
CC Sequence 8 AA;  
CC Query Match 100.0%; Score 66; DB 3; Length 8;  
CC Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC XX  
CC Qy 1 CYWGGYW 8  
CC Db 1 CYWGGYW 8  
CC RESULT 3  
ID ADG94005 standard; peptide: 8 AA.  
XX ADG94005;  
AC AC  
XX DT 11-MAR-2004 (first entry)  
XX DE Immunogenic peptide.  
XX KW Immuno-  
KW genic peptide; multi-specific antibody; polymer conjugate; tumour;  
KW cytostatic; photodynamic therapy.  
XX OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT PH Key  
FT Modified-Site 1  
FT FT /note= "Optionally methylated or Acetylated"  
FT FT  
FT Misc-difference 2  
FT FT /note= "D-form residue"  
FT FT  
FT Misc-difference 3  
FT FT /note= "D-form residue"  
FT FT  
FT Misc-difference 5  
FT FT /note= "Optionally methylated or Acetylated D-form residue"  
FT FT  
FT Misc-difference 7  
FT FT /note= "D-form residue"  
FT FT  
FT XX /note= "D-form residue"  
FT PN US2003026764-A1.  
PN



chosen from bacterial, fungal, parasitic and viral lesion. The infectious disease is caused by a fungus chosen from *Micromsporum*, *Trichophyton*, *Cryptococcus neoformans*, *Sporothrix schenckii*, *Coccidioides immitis*, *Histoplasma capsulatum*, *Blastomyces dermatitidis*, *Candida albicans*. The infectious disease is caused by a virus chosen from HIV, herpes virus, cytomegalovirus, *Sendai* virus, feline leukemia virus, Reo virus, polio virus, human serum parvo-like virus, simian virus 40, respiratory syncytial virus, mouse mammary tumour virus, *Varicella-Zoster* virus, Dengue virus, rubella virus, measles virus, adenovirus, human T-cell leukaemia viruses, Epstein-Barr virus, murine leukaemia virus, mumps virus, vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis virus, wart virus and blue tongue virus. The infectious disease is caused by a bacterium chosen from *Bacillus anthracis*, *Streptococcus agalactiae*, *Legionella pneumophila*, *Streptococcus pyogenes*, *Escherichia coli*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Pneumococcus*, *Haemophilus influenzae* B, *Treponema pallidum*, Lyme disease spirochaetes, *Pseudomonas aeruginosa*, *Mycobacterium leprae*, *Brucella abortus*, *Mycobacterium tuberculosis*, and *Tetanus* toxin. The infectious disease is caused by a protozoa chosen from *Plasmodium falciparum*, *Plasmodium vivax*, *Toxoplasma gondii*, *Trypanosoma rangeli*, *Trypanosoma cruzi*, *Trypanosoma rhodesiense*, *Trypanosoma brucei*, *Schistosoma mansoni*, *Schistosoma japonicum*, *Babesia bovis*, *Elmeria tenella*, *Oncocerca volvulus*, *Leishmania tropica*, *Trichinella spiralis*, *Oncocerca volvulus*, *Theileria parva*, *Taenia hydatigena*, *Tenias ovis*, *Tenias saginata*, *Echinococcus granulosus*, and *Mesocestoides corti*. The infectious disease is caused by a mycoplasma chosen from *Mycoplasma arthritidis*, *M. hyorhinis*, *M. oralis*, *M. arginini*, *Acholeplasma laidlawii*, *M. salivarium* and *M. pneumoniae*. The cancer is preferably chosen from carinoembryonic antigen (CEA)-expressing tumour or a CD20-expressing malignancy. The present sequence represents a peptide used in the method of the invention.

Sequence 8 AA:  
 Query Match Score 100.0%; DB 8; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYWCGGYW 8  
 Db 1 CYWCGGYW 8

RESULT 5  
 AAY66023 standard; peptide: 54 AA.  
 XX AAY66023;  
 XX DR 10-FEB-2000 (first entry)  
 DE Telomerase associated protein TP-1 mutant peptide 6.  
 KW Human; frameshift mutant; T cell response; tumour; treatment; cancer; muttein.  
 KW OS Homo sapiens.  
 OS Synthetic.  
 XX PN WO9958552-A2.  
 XX PD 18-NOV-1999.  
 XX PF 03-MAY-1999; 95WO-NO000143.  
 XX PR 08-MAY-1998; 95WO-00002097.  
 XX PA (HYSD ) NORSK HYDRO AS.

PT New peptides derived from genes with frameshift mutations, used to develop products for the treatment and prophylaxis of cancers.  
 PT XX Claim 13; Page 36; 16pp; English.  
 XX CC Peptides AAY65684-Y66142 are fragments of mutant proteins arising from a frameshift mutation in a gene from a cancer cell. The peptides are characterised in that they: (i) are at least 8 amino acids long and a fragment of a mutant protein arising from a frameshift mutation in a gene of a cancer cell; (ii) consist of at least one amino acid of the mutant part of a protein sequence encoded by the gene; (iii) comprise 0-10 amino acid from the carboxyl terminus of the normal part of the protein sequence preceding the amino terminus of the mutant sequence and may further extend to the carboxyl terminus of the mutant part of the protein as determined by a new stop codon generated by the frameshift mutation; and (iv) induce, either in their full lengths or after processing by an antigen presenting cell (APC), T cell responses. The genes that the peptides are derived from, are characterised as susceptible to frameshift mutation by having a mono nucleotide base repeat sequence of at least 5 residues, or a di-nucleotide base repeat sequence of at least 4 di-nucleotide base units. The peptides are created by the addition or deletion of 1 or 2 nucleotide base residues from the repeat sequence. The novel peptides can elicit T cell responses and toxicity against tumours and cancer cells carrying genes with frameshift mutations. The novel peptides and DNA sequences can be used for the preparation of a composition for the treatment or prophylaxis of cancer.  
 SQ Sequence 54 AA;  
 Query Match Score 72.7%; DB 3; Length 54;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 6  
 ADC33110 standard; protein: 249 AA.  
 XX AC ADC33110;  
 XX DT 18-DEC-2003 (first entry)  
 XX ID ADC33110 standard; protein: 249 AA.  
 XX AC ADC33110;  
 XX DT 18-DEC-2003 (first entry)  
 XX ID ADC33110 standard; protein: 249 AA.  
 XX AC ADC33110;

XX Human novel contig-encoded polypeptide sequence, SEQ ID NO:3192.  
 XX DE Human; diagnostic; drug screening; forensics; gene mapping;  
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
 KW ulcers; osteoporosis; autoimmune disease; cancer;  
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
 KW neuroprotective; antihaemolytic; anticoagulant; thrombolytic; vulnerary;  
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
 KW gene therapy.  
 XX OS Homo sapiens.  
 XX PN WO2003029271-A2.  
 XX PD 10-APR-2003.  
 XX PF 24-SEP-2002; 2002WO-US030474.  
 XX PR 24-SEP-2001; 2001US-0324631P.  
 XX PA (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao Q, Wang J, Wehrman T;  
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
 PI Haley-Vicente D, Dirmancic RT;  
 XX WPI 2000-039064/03.

DR WPI; 2003-371981/35.  
 DR N-PSDB; ADC32343.  
 XX  
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
 PR treating conditions such as neurodegenerative diseases, anemias, platelet  
 PR disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 PR cancer.  
 XX  
 PS Example 2; SEQ ID NO 3192; 1185pp; English.  
 XX  
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-  
 CC ADC30889) and the polypeptides they encode (ADC0890-ADC31860). The  
 CC invention also relates to nucleic acid sequences over 99% identical with  
 CC the novel human cDNAs. The invention additionally encompasses expression  
 CC vectors and host cells comprising a nucleic acid of the invention; the  
 CC recombinant production of a polypeptide of the invention; an antibody  
 CC against a polypeptide of the invention; a method of detecting  
 CC polynucleotides or polypeptides of the invention; and methods of  
 CC identifying a compound which binds to a polypeptide of the invention. The  
 CC invention further discloses methods of preventing, treating or  
 CC ameliorating a medical condition, kits comprising polynucleotide probes  
 CC and/or monoclonal antibodies for carrying out the methods of the  
 CC invention; methods for the identification of compounds that modulate the  
 CC expression or activity of the polynucleotide and/or polypeptide; and 767  
 CC contig sequences corresponding to the cDNA sequences of the invention  
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
 CC -ADC33394). The nucleic acids and polypeptides of the invention are  
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
 CC identification of mutations responsible for genetic disorders or other  
 CC traits, for assessing biodiversity, and in producing many other types of  
 CC data and products dependent on DNA and amino acid sequences. They are  
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
 CC disease and other neurodegenerative diseases, anaemia, platelet  
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 CC cancer. The nucleic acids may also be used as hybridisation probes or  
 CC primers, and in the recombinant production of a protein. The polypeptides  
 CC are also useful in generating antibodies, as molecular weight markers,  
 CC and as food supplements. The present sequence represents a human contig-  
 CC encoded polypeptide sequence used in an example of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).  
 XX  
 Sequence 249 AA;

Query Match 72.7%; Score 48; DB 7; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; OS Unidentified; SEQ ID 33.

QY 1 CYNGCG 6

DB 231 CYNGCG 236

RESULT 7  
 AAB18511 AAB18511 standard; protein; 841 AA.  
 DE H2 homologue of prolyl-tripeptidyl peptidase DPP.  
 XX KW Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;  
 KW gingivitis; periodontitis.  
 XX OS Porphyromonas gingivalis.  
 XX PN WO2000052147-A2.  
 XX PD 08-SEP-2000.

XX PI Han H, Min H, Boone TC;  
 XX  
 PF 03-MAR-2000; 2000WO-US005551.  
 PR 05-MAR-1999; 99US-0123148P.  
 XX (UYGE-) UNIV GEORGIA RES FOUND INC.  
 PA (TRAV/) TRAVIS J.  
 PA (POTE/) POTEMPA J.  
 PA (BANB/) BANBULA A.  
 XX PI Travis J, Potempa J, Banbula A;  
 XX DR 2000-594181/56.  
 XX PT Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful  
 PR for identifying its inhibitor which is useful for protecting an animal  
 PR from a periodontal disease such as gingivitis and periodontitis.  
 XX PS Claim 22; Fig 6; 58pp; English.  
 XX  
 CC The present sequence represents a H2 homologue of a prolyl tripeptidyl-  
 CC peptidase (designated DPP) from *Porphyromonas gingivalis*. The prolyl  
 CC tripeptidyl-peptidase has an amidolytic activity, and cleaves a peptide  
 CC bond in a target polypeptide having at least 4 amino acids. This bond is  
 CC between a proline and an amino acid attached to the alpha-carboxyl group  
 CC end of the proline. The polypeptide is useful for identifying inhibitors.  
 CC These inhibitors are then useful for reducing the growth of bacterium or  
 CC for protecting an animal from a periodontal disease such as gingivitis  
 CC and periodontitis caused by *Porphyromonas gingivalis*  
 XX SQ Sequence 841 AA;

Query	Match	Score	Length
	71.2%	47	841;
Best Local	Similarity	85.7%	
Matches	6; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	Pred. No. 4.2e+02;
Qy	2 YNGCGYW 8		
Db	719 YNGCGYW 725		

RESULT 8  
 ADQ11552  
 ID ADQ11552 standard; peptide; 14 AA.  
 XX  
 AC ADQ11552;  
 AC ADQ11552;  
 AC ADQ11552;  
 XX DT 07-OCT-2004 (first entry)  
 XX DE Myostatin binding peptide, Myostatin-TN8-23, SEQ ID 33.  
 XX KW Muscular; Muscular; Gene Therapy; Antidiabetic; Anorectic; Myostatin;  
 KW myostatin-binding peptide; growth/differentiation factor 8; GDF-8;  
 KW skeletal muscle mass; muscle-wasting disease; muscular dystrophy;  
 KW amyotrophic lateral sclerosis; congestive obstructive pulmonary disease;  
 KW chronic heart failure; cancer; AIDS; renal failure; uremia;  
 KW rheumatoid arthritis; age-related sarcopenia; muscle-wasting;  
 KW spinal chord injury; stroke; bone fracture; aging; diabetes; obesity;  
 KW hyperglycaemia; bone loss; osteoporosis.  
 OS Unidentified.  
 XX PN WO2004058988-A2.  
 XX PD 15-JUL-2004.  
 XX PR 19-DEC-2003; 2003WO-US040781.  
 XX PR 20-DEC-2002; 2002US-0435923P.  
 XX PA (AMGE-) AMGEN INC.  
 XX PI Han H, Min H, Boone TC;

WPI : 2004-525896/50.  
X New binding agent which inhibits myostatin, useful for treating muscular  
X dystrophy, amyotrophic lateral sclerosis, congestive obstructive  
X pulmonary disease, stroke, aging, diabetes, obesity, osteoporosis.  
X Example 1; SEQ ID NO 33; 287pp; English.  
X The present invention relates to novel binding agents comprising at least one peptide capable of binding myostatin and inhibiting its activity. In one embodiment, the binding agent comprises at least one myostatin-binding peptide attached directly or indirectly to at least one vehicle such as a polymer or an FC domain. Myostatin (also known as growth/differentiation factor 8, GDF 8) is a transforming growth-factor-beta (TGF-beta) family member known to be involved in regulation of skeletal muscle mass. The binding agents increase lean muscle mass when administered to animals and decrease fat to muscle ratios. The binding agents are useful for treating muscle-wasting disease, e.g. muscular dystrophy, amyotrophic lateral sclerosis, congestive obstructive pulmonary disease, chronic heart failure, cancer, AIDS, renal failure, uremia, rheumatoid arthritis, age-related sarcopenia, muscle-wasting due to prolonged bedrest, spinal chord injury, stroke, bone fracture, or aging, and myosin-related metabolic disorder, e.g. diabetes, obesity, hyperglycaemia, bone loss, or osteoporosis. The present sequence is one such myostatin-binding peptide.

Sequence 14 AA:	
Query Match	69.7%
Best Local Similarity	71.4%
Matches	5;
Conservative	1;
Mismatches	1;
Indels	0;
gaps	0;

Unidentified. WO2004058988-A2.  
15-JUL-2004.  
19-DEC-2003; 2003WO-US040781.  
20-DEC-2002; 2002US-0435923P.  
(AMGE- ) AMGEN INC.  
Han H, Min H, Boone TC;  
WPI: 2004-525896/50.  
New binding agent which inhibits myostatin, useful for treating muscular

**RESULT 10**  
 ABU33201 standard; protein; 265 AA.  
 XX  
 AC  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DB Protein encoded by Prokaryotic essential gene #18728.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Legionella pneumophila.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-0094893.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C,  
 Wall D, Trawick JD, Carr GJ,  
 XX  
 DR Yamamoto R, Forsyth RA,  
 Haselbeck R, Ohlseni KL,  
 Zyskind JW,  
 Xu HH;  
 WPI; 2003-02926/02.

XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 61125; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: CC (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding antisense nucleic acid; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

XX Sequence 265 AA;

Query Match 69.7%; Score 46; DB 6; Length 265;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CYWGCGYNN 8  
| : | || |:  
Db 65 CFWACGYYF 72

## RESULT 11

ID AAG40353 standard; protein; 233 AA.  
XX  
AC AAG40353;  
XX  
DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 50058.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.

XX PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-012984P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131443P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135359P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138024P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138841P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 22-JUN-1999; 99US-0139839P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140693P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143512P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144066P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144331P.

PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-015925P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-015929P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159310P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159311P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-014484P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159534P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-014518P.	PR	21-OCT-1999;	99US-0160890P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161559P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-014389P.	PR	28-OCT-1999;	99US-0161920P.
PR	04-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161922P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query	1 CY-WGCGYW 8	Score 45.5; DB 3; Length 233;
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity	68.9%; Pred. No. 2.3e+02;	Mismatches 1;
PR	06-AUG-1999;	99US-0147260P.	Matches 7;	Conservative 0;	Gaps 1;
PR	06-AUG-1999;	99US-0147303P.	AAG60115;		
PR	09-AUG-1999;	99US-0147416P.	Db	154 CYRWGCGW 162	
PR	09-AUG-1999;	99US-0147479P.	RESULT 12		
PR	09-AUG-1999;	99US-0147935P.	ID AAG60115 standard; protein; 238 AA.		
PR	10-AUG-1999;	99US-0148170P.	XX		
PR	11-AUG-1999;	99US-0148310P.	AC		
PR	12-AUG-1999;	99US-0148344P.	XX		
PR	13-AUG-1999;	99US-0148565P.	DT	18-OCT-2000 (first entry)	
PR	13-AUG-1999;	99US-0148684P.	XX		
PR	16-AUG-1999;	99US-0149368P.	DS	Arabidopsis thaliana protein fragment SEQ ID NO: 778229.	
PR	17-AUG-1999;	99US-0149175P.	XX		
PR	18-AUG-1999;	99US-0149426P.	KW	Protein identification; signal transduction pathway; metabolic pathway;	
PR	20-AUG-1999;	99US-0149722P.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
PR	20-AUG-1999;	99US-0149723P.	KW	termination sequence.	
PR	20-AUG-1999;	99US-0149929P.	XX		
PR	23-AUG-1999;	99US-0149902P.	OS	Arabidopsis thaliana.	
PR	23-AUG-1999;	99US-0149303P.	XX		
PR	23-AUG-1999;	99US-0150566P.	XX		
PR	26-AUG-1999;	99US-0150884P.	PN	EP1033405-A2.	
PR	27-AUG-1999;	99US-0151065P.	XX		
PR	27-AUG-1999;	99US-0153070P.	PD	06-SEP-2000.	
PR	13-SEP-1999;	99US-0153758P.	XX		
PR	15-SEP-1999;	99US-0151080P.	PP	25-FEB-2000; 2000EP-00301439.	
PR	16-SEP-1999;	99US-015101D.	XX		
PR	20-SEP-1999;	99US-0154030P.	PR	05-MAR-1999;	99US-0121825P.
PR	22-SEP-1999;	99US-0154779P.	PR	09-MAR-1999;	99US-0123180P.
PR	22-SEP-1999;	99US-0155130P.	PR	23-MAR-1999;	99US-0123548P.
PR	24-SEP-1999;	99US-0155650P.	PR	25-MAR-1999;	99US-0126224P.
PR	28-SEP-1999;	99US-0154650P.	PR	29-MAR-1999;	99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999;	99US-0127462P.
PR	04-OCT-1999;	99US-0157110P.	PR	06-APR-1999;	99US-0128224P.
PR	05-OCT-1999;	99US-0157754P.	PR	08-APR-1999;	99US-0128714P.
PR	06-OCT-1999;	99US-0157863P.	PR	16-APR-1999;	99US-0129845P.
PR	07-OCT-1999;	99US-0158028P.	PR	19-APR-1999;	99US-0130077P.
PR	08-OCT-1999;	99US-0158230P.	PR	21-APR-1999;	99US-0130449P.
PR	12-OCT-1999;	99US-0158363P.	PR	23-APR-1999;	99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	23-APR-1999;	99US-0130891P.

PR	28-APR-1999;	99US-0131449P.	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	99US-0145088P.
PR	30-APR-1999;	99US-0112407P.	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	99US-0145088P.
PR	05-MAY-1999;	99US-0132485P.	99US-0145088P.
PR	06-MAY-1999;	99US-0132486P.	99US-0145088P.
PR	07-MAY-1999;	99US-0132487P.	99US-0145192P.
PR	11-MAY-1999;	99US-0132863P.	99US-0145145P.
PR	14-MAY-1999;	99US-0134218P.	99US-0145218P.
PR	14-MAY-1999;	99US-0134219P.	99US-0145218P.
PR	14-MAY-1999;	99US-0134221P.	99US-0145218P.
PR	18-MAY-1999;	99US-0134370P.	99US-0145918P.
PR	19-MAY-1999;	99US-0134768P.	99US-0145919P.
PR	21-MAY-1999;	99US-0135124P.	99US-0145919P.
PR	24-MAY-1999;	99US-0135353P.	99US-0146388P.
PR	04-JUN-1999;	99US-0135629P.	99US-0146388P.
PR	27-MAY-1999;	99US-0136021P.	99US-0146388P.
PR	28-MAY-1999;	99US-0136392P.	99US-0146388P.
PR	01-JUN-1999;	99US-0137222P.	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	99US-0147268P.
PR	04-JUN-1999;	99US-0137502P.	99US-0147268P.
PR	07-JUN-1999;	99US-0137724P.	99US-0147038P.
PR	08-JUN-1999;	99US-0138094P.	99US-0147038P.
PR	10-JUN-1999;	99US-0138540P.	99US-0147204P.
PR	10-JUN-1999;	99US-0138847P.	99US-0147204P.
PR	14-JUN-1999;	99US-0139119P.	99US-0147199P.
PR	16-JUN-1999;	99US-0139452P.	99US-0147268P.
PR	16-JUN-1999;	99US-0139453P.	99US-0147268P.
PR	17-JUN-1999;	99US-0139492P.	99US-0147416P.
PR	18-JUN-1999;	99US-0139454P.	99US-0147493P.
PR	18-JUN-1999;	99US-0139455P.	99US-0147935P.
PR	18-JUN-1999;	99US-0139456P.	99US-0148177P.
PR	18-JUN-1999;	99US-0139457P.	99US-0148319P.
PR	18-JUN-1999;	99US-0139458P.	99US-0148341P.
PR	18-JUN-1999;	99US-0139459P.	99US-0148565P.
PR	18-JUN-1999;	99US-0139460P.	99US-0148684P.
PR	21-JUN-1999;	99US-0139461P.	99US-0149368P.
PR	23-JUN-1999;	99US-0139462P.	99US-0149175P.
PR	23-JUN-1999;	99US-0139463P.	99US-0149428P.
PR	28-JUN-1999;	99US-0139750P.	99US-0149722P.
PR	18-JUN-1999;	99US-0139763P.	99US-0149722P.
PR	29-JUN-1999;	99US-0140991P.	99US-0149904P.
PR	08-JUL-1999;	99US-0142803P.	99US-0149904P.
PR	09-JUL-1999;	99US-0142920P.	99US-0151303P.
PR	01-JUL-1999;	99US-0141842P.	99US-0151438P.
PR	01-JUL-1999;	99US-0142054P.	99US-0151930P.
PR	06-JUL-1999;	99US-0140695P.	99US-0152363P.
PR	08-JUL-1999;	99US-0142390P.	99US-0153070P.
PR	16-JUL-1999;	99US-0144085P.	99US-0153758P.
PR	16-JUL-1999;	99US-0144086P.	99US-015648P.
PR	19-JUL-1999;	99US-0144325P.	99US-0158222P.
PR	19-JUL-1999;	99US-0144331P.	99US-0158389P.
PR	14-JUL-1999;	99US-0144332P.	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	99US-0159638P.

PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160769P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160982P.  
 PR 22-OCT-1999; 99US-0160988P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161932P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 68.9%; Score 45.5; DB 3; Length 238;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+0;  
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CY-WGCGW 8  
 Db 159 CYRWGCGGW 167

---

**RESULT 13**  
**ID AAC05644 standard; protein: 252-AA.**  
**AC AAG05644;**  
**DT 17-OCT-2000 (first entry)**

**DE** Arabidopsis thaliana protein fragment SEQ ID NO: 2124.  
**XX** Protein identification; Signal transduction pathway; metabolic pathway;  
**KW** hybridisation assay; genetic mapping; gene expression; control; promoter;  
**KW** termination sequence.  
**XX** Arabidopsis thaliana.

**OS** Arabidopsis thaliana.  
**XX** EP1033105-A2.  
**XX** PD 06-SEP-2000.  
**PP** 25-FEB-2000; 2000EP-00301439.  
**XX** 25-FEB-1999; 99US-0121823P.  
**PR** 05-MAR-1999; 99US-0123182P.  
**PR** 23-MAR-1999; 99US-0123548P.  
**PR** 23-MAR-1999; 99US-0125788P.  
**PR** 23-MAR-1999; 99US-0126261P.  
**PR** 01-APR-1999; 99US-0126785P.  
**PR** 05-APR-1999; 99US-0127465P.  
**PR** 08-APR-1999; 99US-0128714P.  
**PR** 16-APR-1999; 99US-0129845P.  
**PR** 19-APR-1999; 99US-0130077P.  
**PR** 21-APR-1999; 99US-0130449P.  
**PR** 23-APR-1999; 99US-0130510P.  
**PR** 23-APR-1999; 99US-0130899P.  
**PR** 28-APR-1999; 99US-0131449P.  
**PR** 30-APR-1999; 99US-0132040P.  
**PR** 30-APR-1999; 99US-0132402P.  
**PR** 04-MAY-1999; 99US-0132488P.  
**PR** 05-MAY-1999; 99US-0132489P.  
**PR** 06-MAY-1999; 99US-0132489P.  
**PR** 07-MAY-1999; 99US-0132863P.  
**PR** 11-MAY-1999; 99US-0134256P.  
**PR** 14-MAY-1999; 99US-0134218P.  
**PR** 14-MAY-1999; 99US-0134219P.  
**PR** 14-MAY-1999; 99US-0134211P.  
**PR** 14-MAY-1999; 99US-0134310P.  
**PR** 18-MAY-1999; 99US-0134768P.  
**PR** 19-MAY-1999; 99US-0134911P.  
**PR** 20-MAY-1999; 99US-013524P.  
**PR** 21-MAY-1999; 99US-013533P.  
**PR** 24-MAY-1999; 99US-0135629P.  
**PR** 25-MAY-1999; 99US-0136021P.  
**PR** 27-MAY-1999; 99US-0136322P.  
**PR** 28-MAY-1999; 99US-0136742P.  
**PR** 01-JUN-1999; 99US-0137222P.  
**PR** 03-JUN-1999; 99US-013758P.  
**PR** 04-JUN-1999; 99US-0137502P.  
**PR** 07-JUN-1999; 99US-0137724P.  
**PR** 08-JUN-1999; 99US-013804P.  
**PR** 10-JUN-1999; 99US-0138540P.  
**PR** 10-JUN-1999; 99US-0138847P.  
**PR** 14-JUN-1999; 99US-0139119P.  
**PR** 16-JUN-1999; 99US-0139452P.  
**PR** 16-JUN-1999; 99US-0139453P.  
**PR** 17-JUN-1999; 99US-0139492P.  
**PR** 18-JUN-1999; 99US-0139494P.  
**PR** 18-JUN-1999; 99US-0139455P.  
**PR** 18-JUN-1999; 99US-0139456P.  
**PR** 18-JUN-1999; 99US-0139457P.  
**PR** 18-JUN-1999; 99US-0139458P.  
**PR** 18-JUN-1999; 99US-0139459P.  
**PR** 18-JUN-1999; 99US-0139460P.  
**PR** 18-JUN-1999; 99US-0139462P.  
**PR** 18-JUN-1999; 99US-0139463P.  
**PR** 18-JUN-1999; 99US-0139750P.  
**PR** 18-JUN-1999; 99US-0139753P.  
**PR** 21-JUN-1999; 99US-0140823P.  
**PR** 22-JUN-1999; 99US-0139839P.  
**PR** 23-JUN-1999; 99US-0140353P.  
**PR** 23-JUN-1999; 99US-0140354P.  
**PR** 24-JUN-1999; 99US-0140635P.  
**PR** 28-JUN-1999; 99US-0140911P.  
**PR** 29-JUN-1999; 99US-0140912P.  
**PR** 30-JUN-1999; 99US-0141287P.  
**PR** 01-JUL-1999; 99US-0141342P.  
**PR** 01-JUL-1999; 99US-0142154P.  
**PR** 02-JUL-1999; 99US-0142055P.  
**PR** 06-JUL-1999; 99US-0142300P.  
**PR** 08-JUL-1999; 99US-0142803P.  
**PR** 09-JUL-1999; 99US-0142903P.  
**PR** 12-JUL-1999; 99US-0142977P.  
**PR** 13-JUL-1999; 99US-0143542P.  
**PR** 14-JUL-1999; 99US-0143624P.  
**PR** 15-JUL-1999; 99US-0144005P.  
**PR** 16-JUL-1999; 99US-0144085P.  
**PR** 19-JUL-1999; 99US-0144333P.  
**PR** 19-JUL-1999; 99US-0144334P.  
**PR** 19-JUL-1999; 99US-0144335P.  
**PR** 20-JUL-1999; 99US-0144312P.  
**PR** 20-JUL-1999; 99US-0144632P.  
**PR** 21-JUL-1999; 99US-0144884P.  
**PR** 21-JUL-1999; 99US-0144814P.  
**PR** 21-JUL-1999; 99US-0145086P.  
**PR** 22-JUL-1999; 99US-0145005P.  
**PR** 22-JUL-1999; 99US-0145087P.  
**PR** 22-JUL-1999; 99US-0145089P.  
**PR** 22-JUL-1999; 99US-0145192P.

PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0146390P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147205P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	05-AUG-1999;	99US-0147261P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0147936P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149358P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	23-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150844P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151100P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155599P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0159295P.			
PR	07-OCT-1999;	99US-0158765P.			
PR	14-OCT-1999;	99US-0158929P.			
PR	14-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158362P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	18-OCT-1999;	99US-0159295P.			
PR	21-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			

		Query Match	68.9%;	Score 45.5;	Length 252;
		Best Local Similarity	77.8%;	Pred. No. 2.4e+02;	
		Matches	7;	Mismatches 0;	
		Conservative	0;	Indels 1;	Gaps 1;
Qy		1 CY-WGCGGW 8			
		Db	173 CYWGCGGW 181		
<b>RESULT 14</b>					
	AAG40352	AAAG40352 standard; protein; 252 AA.			
	ID	AAG40352;			
	XX	XX			
	AC	AC			
	XX	XX			
	DT	18-OCT-2000 (first entry)			
	XX	XX			
	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 50057.			
	XX	KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; gene expression control; promoter; termination sequence.			
	XX	Arabidopsis thaliana.			
	OS	OS			
	XX	XX			
	PN	EP1033405-A2.			
	XX	XX			
	PD	06-SEP-2000.			
	XX	XX			
	PF	25-FEB-2000; 2000EP-00301439.			
	XX	XX			
	PR	25-FEB-1999; 99US-0121825P.			
	PR	05-MAR-1999; 99US-0123180P.			
	PR	09-MAR-1999; 99US-0123548P.			
	PR	23-MAR-1999; 99US-0125798P.			
	PR	25-MAR-1999; 99US-0126264P.			
	PR	29-MAR-1999; 99US-0126795P.			
	PR	01-APR-1999; 99US-0127462P.			
	PR	05-MAR-1999; 99US-0128244P.			
	PR	08-APR-1999; 99US-0128714P.			
	PR	23-APR-1999; 99US-0130077P.			
	PR	16-APR-1999; 99US-0130078P.			
	PR	19-APR-1999; 99US-0130079P.			
	PR	21-APR-1999; 99US-0130449P.			
	PR	04-MAY-1999; 99US-0130510P.			
	PR	05-MAY-1999; 99US-0130891P.			
	PR	23-APR-1999; 99US-0132485P.			
	PR	06-MAY-1999; 99US-0132486P.			
	PR	08-APR-1999; 99US-0132487P.			
	PR	14-MAY-1999; 99US-0134219P.			
	PR	14-MAY-1999; 99US-0134221P.			
	PR	14-MAY-1999; 99US-0134370P.			
	PR	18-MAY-1999; 99US-0134768P.			

PR	19-MAY-1999;	99US-0134941P.	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	99US-0146386P.
PR	21-MAY-1999;	99US-013535P.	99US-0146388P.
PR	24-MAY-1999;	99US-013562P.	99US-0146389P.
PR	25-MAY-1999;	99US-013602P.	99US-0147038P.
PR	27-MAY-1999;	99US-013639P.	99US-014704P.
PR	28-MAY-1999;	99US-01378P.	99US-0147302P.
PR	01-JUN-1999;	99US-013722P.	99US-0147192P.
PR	03-JUN-1999;	99US-013752P.	99US-0147260P.
PR	04-JUN-1999;	99US-013750P.	99US-0147303P.
PR	07-JUN-1999;	99US-013772P.	99US-0147416P.
PR	08-JUN-1999;	99US-013809P.	99US-0147493P.
PR	10-JUN-1999;	99US-013154P.	99US-014735P.
PR	18-JUN-1999;	99US-013884P.	99US-0149171P.
PR	14-JUN-1999;	99US-013911P.	99US-0148319P.
PR	16-JUN-1999;	99US-013945P.	99US-0148341P.
PR	17-JUN-1999;	99US-013945P.	99US-0148565P.
PR	18-JUN-1999;	99US-013949P.	99US-014884P.
PR	18-JUN-1999;	99US-013945P.	99US-0149368P.
PR	18-JUN-1999;	99US-013945P.	99US-014910P.
PR	18-JUN-1999;	99US-013945P.	99US-0149226P.
PR	18-JUN-1999;	99US-013945P.	99US-0149228P.
PR	18-JUN-1999;	99US-013945P.	99US-014923P.
PR	18-JUN-1999;	99US-013945P.	99US-0149239P.
PR	18-JUN-1999;	99US-013946P.	99US-0149002P.
PR	18-JUN-1999;	99US-013946P.	99US-0149002P.
PR	18-JUN-1999;	99US-013946P.	99US-0149002P.
PR	18-JUN-1999;	99US-013946P.	99US-0149002P.
PR	18-JUN-1999;	99US-013946P.	99US-0149002P.
PR	18-JUN-1999;	99US-013975P.	99US-015085P.
PR	18-JUN-1999;	99US-013976P.	99US-0151066P.
PR	21-JUN-1999;	99US-013981P.	99US-0151080P.
PR	29-JUN-1999;	99US-013981P.	99US-015103P.
PR	30-JUN-1999;	99US-013989P.	99US-015138P.
PR	23-JUN-1999;	99US-014035P.	99US-0151330P.
PR	23-JUN-1999;	99US-014035P.	99US-0151363P.
PR	24-JUN-1999;	99US-014065P.	99US-0153070P.
PR	28-JUN-1999;	99US-014082P.	99US-015370P.
PR	29-JUN-1999;	99US-014099P.	99US-0153758P.
PR	09-JUL-1999;	99US-014128P.	99US-0154018P.
PR	01-JUL-1999;	99US-014184P.	99US-0154039P.
PR	01-JUL-1999;	99US-014215P.	99US-0154779P.
PR	02-JUL-1999;	99US-014205P.	99US-0155139P.
PR	06-JUL-1999;	99US-014239P.	99US-0155466P.
PR	08-JUL-1999;	99US-014280P.	99US-0155639P.
PR	09-JUL-1999;	99US-014290P.	99US-0156558P.
PR	12-JUL-1999;	99US-014297P.	99US-015656P.
PR	13-JUL-1999;	99US-014354P.	99US-0157117P.
PR	14-JUL-1999;	99US-014362P.	99US-015753P.
PR	15-JUL-1999;	99US-014405P.	99US-015785P.
PR	16-JUL-1999;	99US-0144085P.	99US-0158029P.
PR	19-JUL-1999;	99US-0144334P.	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	99US-0159330P.
PR	19-JUL-1999;	99US-0144335P.	99US-0159331P.
PR	20-JUL-1999;	99US-0144335P.	99US-0160768P.
PR	22-JUL-1999;	99US-0144335P.	99US-0160770P.
PR	22-JUL-1999;	99US-0144332P.	99US-0160814P.
PR	21-JUL-1999;	99US-0144333P.	99US-0160815P.
PR	21-JUL-1999;	99US-0145086P.	99US-016090P.
PR	23-JUL-1999;	99US-0145145P.	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	99US-0160989P.
PR	23-JUL-1999;	99US-0145224P.	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	99US-0161359P.

PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.  
  
 Query Match 68 9%; Score 45.5; DB 3; Length 252;  
 Best Local Similarity 77.8%; Pred. No. 2.4e-02;  
 Matches 7; Conservative 0; Mismatches 1; Gaps 1;  
 AC 1 CY-WCCGYW 8  
 DB 173 CYRNCGW 181  
  
**RESULT 15**  
 AG3-8268 standard; protein; 252 AA.  
 ID AAG38268  
 XX AAG38268;  
 AC AC  
 XX DT 18-OCT-2000 (first entry)  
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47184.  
 XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX OS Arabidopsis thaliana.  
 XX PN EP1033405-A2.  
 XX PD 06-SEP-2000.  
 XX PF 25-FEB-2000; 2000EP-00301439.  
 XX PR 25-FEB-1999; 99US-0121825P.  
 PR 09-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-012578P.  
 PR 29-MAR-1999; 99US-0128264P.  
 PR 01-APR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127452P.  
 PR 08-APR-1999; 99US-0128234P.  
 PR 16-APR-1999; 99US-0128714P.  
 PR 21-APR-1999; 99US-0130077P.  
 PR 23-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 14-MAY-1999; 99US-0132419P.  
 PR 06-MAY-1999; 99US-01311449P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 18-MAY-1999; 99US-0132843P.  
 PR 14-MAY-1999; 99US-0134256P.  
 PR 20-MAY-1999; 99US-0134218P.  
 PR 21-MAY-1999; 99US-0134353P.  
 PR 24-MAY-1999; 99US-015629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139761P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 23-JUL-1999; 99US-0145086P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145520P.  
 PR 26-JUL-1999; 99US-0145087P.  
 PR 26-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145145P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147030P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.



Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:52:14 ; Search time 58:6047 Seconds  
 (without alignments)  
 45.371 Million cell updates/sec

Title: SEQ5

Perfect score: 66

Sequence: 1 cywggyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:  
 1: /cggn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
 2: /cggn2\_6/ptodata/1/pubpaa/PCTUS\_PUB.pep.\*  
 3: /cggn2\_6/ptodata/1/pubpaa/US06\_NEWPUB.pep.\*  
 4: /cggn2\_6/ptodata/1/pubpaa/US07\_NEWPUB.pep.\*  
 5: /cggn2\_6/ptodata/1/pubpaa/PUBCOMB.pep.\*  
 6: /cggn2\_6/ptodata/1/pubpaa/NEW\_PUB.pep.\*  
 7: /cggn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
 8: /cggn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
 9: /cggn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
 10: /cggn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
 11: /cggn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
 12: /cggn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
 13: /cggn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
 14: /cggn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
 15: /cggn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
 16: /cggn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
 17: /cggn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
 18: /cggn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
 19: /cggn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
 20: /cggn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1  
US-10-776-224-340

; Sequence 340, Application US-10776224  
 ; Publication No. US20050074849A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gaudernack, Gustav  
 ; APPLICANT: Bricksen, Jon Amund  
 ; APPLICANT: Moller, Mona  
 ; APPLICANT: Saeterdal, Ingvil  
 ; TITLE OF INVENTION: Peptides  
 ; FILE REFERENCE: 01702\_4015LO  
 ; CURRENT APPLICATION NUMBER: US-10/776,224  
 ; CURRENT FILING DATE: 2004-02-12  
 ; PRIOR APPLICATION NUMBER: US 09/674,973  
 ; PRIOR FILING DATE: 2001-06-04  
 ; NUMBER OF SEQ ID NOS: 459  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO: 340  
 ; LENGTH: 54  
 ; TYPE: PPT  
 ; ORGANISM: Homo sapiens

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	72.7	54	17 US-10-776-224-340	Sequence 340, APP
2	46	69.7	14	16 US-10-742-379-33	Sequence 33, APP
3	46	69.7	50	16 US-10-742-379-136	Sequence 136, APP
4	46	69.7	265	15 US-10-282-122A-61125	Sequence 61125, APP
5	46	69.7	283	16 US-10-757-701-42031	Sequence 42031, APP
6	45.5	68.9	72	16 US-10-437-963-190474	Sequence 190474,
7	45.5	68.9	292	15 US-10-425-114-69268	Sequence 214537,
8	45.5	68.9	444	15 US-10-369-493-21261	Sequence 69268,
9	45	68.2	444	15 US-10-369-493-21261	Sequence 21261, APP
10	45	68.2	404	16 US-10-437-963-155759	Sequence 155759,
11	44	66.7	304	16 US-10-424-599-238640	Sequence 238640,
12	42	63.6	46	15 US-10-106-698-6232	Sequence 6232, APP
13	42	63.6	345	14 US-10-106-698-6232	Sequence 6232, APP

RESULT 2  
US-10-742-379-33

; Sequence 1 CRYGGYW 8  
 ; Sequence 5 CEWGGSW 12

RESULT 2  
US-10-742-379-33  
 ; Sequence 33, Application US/10742379

Publication No. US20040181033A1  
 GENERAL INFORMATION:  
 APPLICANT: Han, HO  
 APPLICANT: Min, Hosung  
 APPLICANT: Boone, Thomas Charles  
 TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN  
 FILE REFERENCE: A-828 (US)  
 CURRENT FILING DATE: 2003-12-19  
 PRIOR APPLICATION NUMBER: US 60/435, 923  
 PRIOR FILING DATE: 2002-12-20  
 NUMBER OF SEQ ID NOS: 634  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 33  
 LENGTH: 14  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Myostatin Binding Peptide  
 US-10-742-379-33

Query Match 69.7%; Score 46; DB 16; Length 14;  
 Best Local Similarity 71.4%; Pred. No. 13;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 SEQ ID NO: 136

RESULT 3  
 US-10-742-379-136  
 Sequence 136, Application US/107423379  
 GENERAL INFORMATION:  
 APPLICANT: Han, HO  
 APPLICANT: Min, Hosung  
 APPLICANT: Boone, Thomas Charles  
 TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN  
 FILE REFERENCE: A-828 (US)  
 CURRENT FILING DATE: 2003-12-19  
 PRIOR APPLICATION NUMBER: US 60/435, 923  
 PRIOR FILING DATE: 2002-12-20  
 NUMBER OF SEQ ID NOS: 634  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 136  
 LENGTH: 50  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Myostatin Binding Peptide  
 US-10-742-379-133

Query Match 69.7%; Score 46; DB 16; Length 50;  
 Best Local Similarity 71.4%; Pred. No. 32;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 SEQ ID NO: 136

RESULT 4  
 US-10-282-122A-61125  
 Sequence 61125, Application US/10282122A  
 GENERAL INFORMATION:  
 APPLICANT: Zanudio, Carlos  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohisen, Kari

Query Match 69.7%; Score 46; DB 16; Length 283;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;

APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forgyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA-034A  
 CURRENT APPLICATION NUMBER: US/10/282, 122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191, 078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/205, 848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207, 727  
 PRIOR FILING DATE: 2000-05-16  
 PRIOR APPLICATION NUMBER: 60/230, 335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230, 347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242, 578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253, 625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257, 931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267, 636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269, 308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 61125  
 LENGTH: 265  
 TYPE: PRT  
 ORGANISM: Legionella pneumophila  
 US-10-282-122A-61125

Query Match 69.7%; Score 46; DB 15; Length 265;  
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CYWGCGW 8  
 Db 65 CFWACGYF 72

RESULT 5  
 US-10-767-701-42031  
 Sequence 42031, Application US/10767701  
 GENERAL INFORMATION:  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 FILE REFERENCE: 38-21 (53535)B  
 CURRENT APPLICATION NUMBER: US/10/767,701  
 CURRENT FILING DATE: 2004-01-29  
 NUMBER OF SEQ ID NOS: 63128  
 SEQ ID NO: 42031  
 LENGTH: 283  
 TYPE: PRT  
 ORGANISM: Sorghum bicolor  
 FEATURE:  
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-C13148\_1.pep  
 US-10-767-701-42031

Query Match 69.7%; Score 46; DB 16; Length 283;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;

```

seq5.rapb

Matches   6;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
Qy        1 CYWGGGY 7
Db        1 CKWGCGY 7

RESULT 6
US-10-437-963-190474
; Sequence 190474, Application US/10437963
; Publication No. US20040123343A1
GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 190474
LENGTH: 72
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_868886C.1.pep
US-10-437-963-190474

Query Match 68.9%; Score 45.5; DB 16; Length 72;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
Number of SEQ ID NOS: 204966
Db      1 CYWGC---GYW 8
Db      55 CRWGCCHRGYW 65

RESULT 7
US-10-424-599-214537
; Sequence 214537, Application US/10424599
; Publication No. US20040031072A1
GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 214537
LENGTH: 217
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3347_35754C.1.pep
US-10-424-599-214537

Query Match 68.9%; Score 45.5; DB 15; Length 217;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Number of SEQ ID NOS: 285684
Db      1 CY-WGCGYW 8
Db      138 CYRWGCGGW 146

RESULT 8
US-10-425-114-69268
; Sequence 6928, Application US/10425114
; Publication No. US2004034888A1
GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 69268
LENGTH: 292
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: JC-GMFL02220141E06_FLI.pep
US-10-425-114-69268

Query Match 68.9%; Score 45.5; DB 15; Length 292;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Number of SEQ ID NOS: 73128
Qy      1 CY-WGCGYW 8
Db      213 CYRWGCGGW 221

RESULT 9
US-10-169-493-21261
; Sequence 21261, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-03-28
PRIORITY APPLICATION NUMBER: US 60/360,039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 21261
LENGTH: 444
TYPE: PRT
ORGANISM: Xenorhabdus nematophilus
US-10-169-493-21261

Query Match 68.2%; Score 45; DB 15; Length 444;
Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Number of SEQ ID NOS: 47374
Qy      2 YNGC---GYW 8
Db      84 YNGCLTAWGYW 94

RESULT 10
US-10-369-493-463
; Sequence 463, Application US/10369493
; Publication No. US20030233675A1

```

GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 CURRENT FILING DATE: 2003-05-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 SEQ ID NO: 463  
 LENGTH: 445  
 TYPE: PRT  
 ORGANISM: Xenorhabdus nematophilus  
 US-10-369-493-463

Query Match 68.2%; Score 45; DB 15; Length 445;  
 Best Local Similarity 63.6%; Pred. No. 2.2e+02; Indels 4; Gaps 1;  
 Matches 7; Conservative 0; Mismatches 0;

Qy 2 YNGC---GYW 8  
 Db 80 YNGGTTAGCYW 90

RESULT 11  
 US-10-437-963-155759  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping  
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Title of Invention: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(52221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 SEQ ID NO 155759  
 NUMBER OF SEQ ID NOS: 204966  
 TYPE: PRT  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_55493C.1.pep

Query Match 66.7%; Score 44; DB 16; Length 304;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 2; Mismatches 1;

Qy 1 CYNGCGYW 8  
 Db 275 CFYCCGNW 282

RESULT 12  
 US-10-424-599-238640  
 Publication No. US2004001072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua

GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285584  
 SEQ ID NO 285584  
 LENGTH: 46  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_57517C.1.pep  
 US-10-424-599-238640

Query Match 63.6%; Score 42; DB 15; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Mismatches 0; Gaps 0;  
 Matches 5; Conservative 0;

Qy 1 CYWGC 5  
 Db 38 CYWGC 42

RESULT 13  
 US-10-106-698-6232  
 Sequence 6232, Application US/10106698  
 Publication No. US20030109690A1  
 GENERAL INFORMATION:  
 APPLICANT: Ruben et al.  
 TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
 FILE REFERENCE: PA005P1  
 CURRENT APPLICATION NUMBER: US/10/106,698  
 CURRENT FILING DATE: 2002-03-27  
 PRIOR APPLICATION NUMBER: PCT/US00/26524  
 PRIOR FILING DATE: 2000-09-28  
 PRIOR APPLICATION NUMBER: US 60/157,137  
 PRIOR FILING DATE: 1999-09-29  
 PRIOR APPLICATION NUMBER: US 60/163,280  
 PRIOR FILING DATE: 1999-11-03  
 NUMBER OF SEQ ID NOS: 8564  
 SOFTWARE: PatentIn Ver. 3.0  
 SEQ ID NO 6232  
 LENGTH: 345  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-106-698-6232

Query Match 63.6%; Score 42; DB 14; Length 345;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+02; Indels 0; Mismatches 0; Gaps 0;  
 Matches 5; Conservative 0;

Qy 1 CYWGC 5  
 Db 111 CYWGC 115

RESULT 14  
 US-10-704-365-2  
 Sequence 2, Application US/10704365  
 Publication No. US20040131605A1  
 GENERAL INFORMATION:  
 APPLICANT: Qin et al.  
 TITLE OF INVENTION: PRECURSOR OF N-ACETYLGLACTOOSAMINE-4-SULFATASE, METHODS OF TREATING  
 FILE REFERENCE: 30610/30011A  
 CURRENT APPLICATION NUMBER: US/10/704,365  
 CURRENT FILING DATE: 2003-11-07  
 PRIOR APPLICATION NUMBER: US 10/290,508  
 PRIOR FILING DATE: 2002-11-07  
 PRIOR APPLICATION NUMBER: US 10/317,249  
 PRIOR FILING DATE: 2002-12-10  
 PRIOR APPLICATION NUMBER: To be assigned

```

; PRIOR FILING DATE: 2003-09-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 2
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-704-365-2

Query Match          63.6%; Score 42; DB 16; Length 533;
Best Local Similarity 85.7%; Pred. No. 6.2e+02;
Matches   6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          2 YWGCGYW 8
Db        444 YPGCGYW 450

RESULT 15
US-10-755-889-308
; Sequence 308, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-κB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 308
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-308

Query Match          63.6%; Score 42; DB 16; Length 533;
Best Local Similarity 85.7%; Pred. No. 6.2e+02;
Matches   6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          2 YWGCGYW 8
Db        444 YPGCGYW 450

```

Search completed: April 18, 2005, 20:29:28

Job time : 59.6047 secs



GenCore version 5.1.6	28	39	59.1	409	4	US-10-117-005-4
Copyright (c) 1993 - 2005 Compugen Ltd.	29	39	59.1	749	4	US-09-489-039A-12279
OM protein - protein search, using sw model	30	38.5	58.3	469	2	US-08-484-126-1
run on: April 18, 2005, 19:38:48 ; Search time 22.3256 Seconds (without alignments) 26.749 Million cell updates/sec	31	38.5	58.3	469	4	Sequence 1, Appli
title: SEQ5	32	38.5	58.3	632	2	Sequence 2, Appli
perfect score: 66	33	38.5	58.3	632	4	Sequence 3, Appli
sequence: 1 cywgcgwy 8	34	38.5	58.3	665	4	Sequence 4, Appli
scoring table: BLOSUM62	35	38.5	58.3	665	4	Sequence 5, Appli
Gapext 10.0 , Gapext 0.5	36	38.5	58.3	1312	4	Sequence 6, Appli
searched: 51355 seqs, 74649064 residues	37	38	57.6	58	4	Sequence 7, Appli
Total number of hits satisfying chosen parameters: 513545	38	38	57.6	87	4	Sequence 8, Appli
	39	38	57.6	111	4	Sequence 9, Appli
	40	38	57.6	164	4	Sequence 10, Appli
	41	38	57.6	383	4	Sequence 11, Appli
	42	37.5	56.8	55	4	Sequence 12, Appli
	43	37.5	56.8	339	4	Sequence 13, Appli
	44	37	56.1	84	4	Sequence 14, Appli
	45	37	56.1	154	4	Sequence 15, Appli
						Sequence 16, Appli
						Sequence 17, Appli
						Sequence 18, Appli
						Sequence 19, Appli
						Sequence 20, Appli
						Sequence 21, Appli
						Sequence 22, Appli
						Sequence 23, Appli
						Sequence 24, Appli
						Sequence 25, Appli
						Sequence 26, Appli
						Sequence 27, Appli
						Sequence 28, Appli
						Sequence 29, Appli
						Sequence 30, Appli
						Sequence 31, Appli
						Sequence 32, Appli
						Sequence 33, Appli
						Sequence 34, Appli
						Sequence 35, Appli
						Sequence 36, Appli
						Sequence 37, Appli
						Sequence 38, Appli
						Sequence 39, Appli
						Sequence 40, Appli
						Sequence 41, Appli
						Sequence 42, Appli
						Sequence 43, Appli
						Sequence 44, Appli
						Sequence 45, Appli
						Sequence 46, Appli
						Sequence 47, Appli
						Sequence 48, Appli
						Sequence 49, Appli
						Sequence 50, Appli
						Sequence 51, Appli
						Sequence 52, Appli
						Sequence 53, Appli
						Sequence 54, Appli
						Sequence 55, Appli
						Sequence 56, Appli
						Sequence 57, Appli
						Sequence 58, Appli
						Sequence 59, Appli
						Sequence 60, Appli
						Sequence 61, Appli
						Sequence 62, Appli
						Sequence 63, Appli
						Sequence 64, Appli
						Sequence 65, Appli
						Sequence 66, Appli
						Sequence 67, Appli
						Sequence 68, Appli
						Sequence 69, Appli
						Sequence 70, Appli
						Sequence 71, Appli
						Sequence 72, Appli
						Sequence 73, Appli
						Sequence 74, Appli
						Sequence 75, Appli
						Sequence 76, Appli
						Sequence 77, Appli
						Sequence 78, Appli
						Sequence 79, Appli
						Sequence 80, Appli
						Sequence 81, Appli
						Sequence 82, Appli
						Sequence 83, Appli
						Sequence 84, Appli
						Sequence 85, Appli
						Sequence 86, Appli
						Sequence 87, Appli
						Sequence 88, Appli
						Sequence 89, Appli
						Sequence 90, Appli
						Sequence 91, Appli
						Sequence 92, Appli
						Sequence 93, Appli
						Sequence 94, Appli
						Sequence 95, Appli
						Sequence 96, Appli
						Sequence 97, Appli
						Sequence 98, Appli
						Sequence 99, Appli
						Sequence 100, Appli
						Sequence 101, Appli
						Sequence 102, Appli
						Sequence 103, Appli
						Sequence 104, Appli
						Sequence 105, Appli
						Sequence 106, Appli
						Sequence 107, Appli
						Sequence 108, Appli
						Sequence 109, Appli
						Sequence 110, Appli
						Sequence 111, Appli
						Sequence 112, Appli
						Sequence 113, Appli
						Sequence 114, Appli
						Sequence 115, Appli
						Sequence 116, Appli
						Sequence 117, Appli
						Sequence 118, Appli
						Sequence 119, Appli
						Sequence 120, Appli
						Sequence 121, Appli
						Sequence 122, Appli
						Sequence 123, Appli
						Sequence 124, Appli
						Sequence 125, Appli
						Sequence 126, Appli
						Sequence 127, Appli
						Sequence 128, Appli
						Sequence 129, Appli
						Sequence 130, Appli
						Sequence 131, Appli
						Sequence 132, Appli
						Sequence 133, Appli
						Sequence 134, Appli
						Sequence 135, Appli
						Sequence 136, Appli
						Sequence 137, Appli
						Sequence 138, Appli
						Sequence 139, Appli
						Sequence 140, Appli
						Sequence 141, Appli
						Sequence 142, Appli
						Sequence 143, Appli
						Sequence 144, Appli
						Sequence 145, Appli
						Sequence 146, Appli
						Sequence 147, Appli
						Sequence 148, Appli
						Sequence 149, Appli
						Sequence 150, Appli
						Sequence 151, Appli
						Sequence 152, Appli
						Sequence 153, Appli
						Sequence 154, Appli
						Sequence 155, Appli
						Sequence 156, Appli
						Sequence 157, Appli
						Sequence 158, Appli
						Sequence 159, Appli
						Sequence 160, Appli
						Sequence 161, Appli
						Sequence 162, Appli
						Sequence 163, Appli
						Sequence 164, Appli
						Sequence 165, Appli
						Sequence 166, Appli
						Sequence 167, Appli
						Sequence 168, Appli
						Sequence 169, Appli
						Sequence 170, Appli
						Sequence 171, Appli
						Sequence 172, Appli
						Sequence 173, Appli
						Sequence 174, Appli
						Sequence 175, Appli
						Sequence 176, Appli
						Sequence 177, Appli
						Sequence 178, Appli
						Sequence 179, Appli
						Sequence 180, Appli
						Sequence 181, Appli
						Sequence 182, Appli
						Sequence 183, Appli
						Sequence 184, Appli
						Sequence 185, Appli
						Sequence 186, Appli
						Sequence 187, Appli
						Sequence 188, Appli
						Sequence 189, Appli
						Sequence 190, Appli
						Sequence 191, Appli
						Sequence 192, Appli
						Sequence 193, Appli
						Sequence 194, Appli
						Sequence 195, Appli
						Sequence 196, Appli
						Sequence 197, Appli
						Sequence 198, Appli
						Sequence 199, Appli
						Sequence 200, Appli
						Sequence 201, Appli
						Sequence 202, Appli
						Sequence 203, Appli
						Sequence 204, Appli
						Sequence 205, Appli
						Sequence 206, Appli
						Sequence 207, Appli
						Sequence 208, Appli
						Sequence 209, Appli
						Sequence 210, Appli
						Sequence 211, Appli
						Sequence 212, Appli
						Sequence 213, Appli
						Sequence 214, Appli
						Sequence 215, Appli
						Sequence 216, Appli
						Sequence 217, Appli
						Sequence 218, Appli
						Sequence 219, Appli
						Sequence 220, Appli
						Sequence 221, Appli
						Sequence 222, Appli
						Sequence 223, Appli
						Sequence 224, Appli
						Sequence 225, Appli
						Sequence 226, Appli
						Sequence 227, Appli
						Sequence 228, Appli
						Sequence 229, Appli
						Sequence 230, Appli
						Sequence 231, Appli
						Sequence 232, Appli
						Sequence 233, Appli
						Sequence 234, Appli
						Sequence 235, Appli
						Sequence 236, Appli
						Sequence 237, Appli
						Sequence 238, Appli
						Sequence 239, Appli
						Sequence 240, Appli
						Sequence 241, Appli
						Sequence 242, Appli
						Sequence 243, Appli
						Sequence 244, Appli
						Sequence 245, Appli
						Sequence 246, Appli
						Sequence 247, Appli
						Sequence 248, Appli
						Sequence 249, Appli
						Sequence 250, Appli
						Sequence 251, Appli
						Sequence 252, Appli
						Sequence 253, Appli
						Sequence 254, Appli
						Sequence 255, Appli
						Sequence 256, Appli
						Sequence 257, Appli
						Sequence 258, Appli
						Sequence 259, Appli
						Sequence 260, Appli
						Sequence 261, Appli
						Sequence 262, Appli
						Sequence 263, Appli
						Sequence 264, Appli
						Sequence 265, Appli
						Sequence 266, Appli
						Sequence 267, Appli
						Sequence 268, Appli
						Sequence 269, Appli
						Sequence 270, Appli
						Sequence 271, Appli
						Sequence 272, Appli
						Sequence 273, Appli
						Sequence 274, Appli
						Sequence 275, Appli
						Sequence 276, Appli
						Sequence 277, Appli
						Sequence 278, Appli
						Sequence 279, Appli
						Sequence 280, Appli
						Sequence 281, Appli
						Sequence 282, Appl

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/445,586  
 PRIORITY DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/111,887  
 FILING DATE: 26-AUG-1993  
 APPLICATION NUMBER: JP 230030/92  
 FILING DATE: 28-AUG-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 324034/92  
 FILING DATE: 03-DEC-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Forman, David S.  
 REFERENCE/DOCKET NUMBER: 33-694  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4000  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 533 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 US-08-445-586-10

Query Match Best Local Similarity 63.6%; Score 42; DB 1; Length 533;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2.YWGCGYW 8  
 Db 444 YPGGGYW 450

RESULT 3 US-08-445-586-13  
 Sequence 13, Application US/08484493  
 Patent No. 5728381  
 GENERAL INFORMATION:  
 APPLICANT: Wilson, Peter J  
 APPLICANT: Morris, Charles P  
 APPLICANT: Anson, Donald S  
 APPLICANT: Occhiodoro, Teresa  
 APPLICANT: Bielicki, Julie  
 APPLICANT: Clements, Peter R  
 APPLICANT: Hopwood, John J  
 TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSEE: Scully, Murphy & Presser  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 11530  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484,494  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 991,973  
 FILING DATE: 17-DEC-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DiGiglio, Frank S  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8416Z  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4343  
 TELEFAX: 516-742-4366  
 TELEX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 533 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-484-494-13

Query Match 63.6%; Score 42; DB 1; Length 533;  
Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGGYW 8  
Db 444 YPGCCYW 450

RESULT 5  
US-08-345-212-13  
Sequence 11, Application US/08345211

GENERAL INFORMATION:  
Patent No. 5932211  
APPLICANT: Wilson, Peter J  
APPLICANT: Morris, Charles P  
APPLICANT: Anson, Donald S  
APPLICANT: Occhiodoro, Teresa  
APPLICANT: Bielicki, Julie  
APPLICANT: Clements, Peter R  
APPLICANT: Hopwood, John J  
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF IDURONATE 2-SULFATASE  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,003  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/991,973  
FILING DATE: 17-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DiGilio, Frank S  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8416Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 533 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-249-003-13

Query Match 63.6%; Score 42; DB 3; Length 533;  
Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGGYW 8  
Db 444 YPGCCYW 450

RESULT 7  
US-09-085-844-13  
Sequence 13, Application US/09685844  
Patent No. 6541254  
GENERAL INFORMATION:  
APPLICANT: Wilson, Peter J  
APPLICANT: Morris, Charles P  
Anson, Donald S  
Occhiodoro, Teresa  
Bielicki, Julie  
Clements, Peter R  
Hopwood, John J  
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF IDURONATE 2-SULFATASE

Query Match 63.6%; Score 42; DB 2; Length 533;  
Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGGYW 8  
Db 444 YPGCCYW 450

RESULT 6

NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Scort, Murphy & Presser  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 11530

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/469,260A  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/424,550  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: FORMSKI, PRISCILLA E.  
 REGISTRATION NUMBER: 33,207  
 REFERENCE/DOCKET NUMBER: 5527.PC.01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 708-937-6365  
 TELEFAX: 708-938-2623  
 INFORMATION FOR SEQ ID NO: 387:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2972 amino acids  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DiGilio, Frank S  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8416Z  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4343  
 TELEFAX: 516-742-4366  
 TELEX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 533 amino acids  
 STRANDBNESS: single  
 TOPOLogy: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
 US-09-685-844-13

Query Match 63.6% Score 42; DB 4; Length 533;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;

Qy	2 YWGCGYW 8
Db	444 YWGCGYW 450

RESULT 9  
 US-08-488-446-387  
 Sequence 387, Application US/08488446  
 Patent No. 6558898

GENERAL INFORMATION:  
 APPLICANT: JOHN N. SIMONS  
 APPLICANT: TAMI J. PILOT-MATIAS  
 APPLICANT: GEORGE J. DAWSON  
 APPLICANT: GEORGE G. SCHLAUDER  
 APPLICANT: SURESH M. DESAI  
 APPLICANT: THOMAS P. LEARY  
 APPLICANT: ANTHONY SCOTT MUERHOF  
 APPLICANT: JAMES C. ERKER  
 APPLICANT: SHERI L. BULJK  
 APPLICANT: ISA K. MUSHAHWAR  
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 NUMBER OF SEQUENCES: 716  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ABBOTT LABORATORIES  
 STREET: 100 ABBOTT PARK ROAD  
 CITY: ABBOTT PARK  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/488,446  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/424,550  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:

RESULT 8  
 US-08-469-260A-387  
 Sequence 387, Application US/08469260A  
 Patent No. 6451578

GENERAL INFORMATION:  
 APPLICANT: JOHN N. SIMONS  
 APPLICANT: TAMI J. PILOT-MATIAS  
 APPLICANT: GEORGE J. DAWSON  
 APPLICANT: GEORGE G. SCHLAUDER  
 APPLICANT: SURESH M. DESAI  
 APPLICANT: THOMAS P. LEARY  
 APPLICANT: ANTHONY SCOTT MUERHOF  
 APPLICANT: JAMES C. ERKER  
 APPLICANT: SHERI L. BULJK  
 APPLICANT: ISA K. MUSHAHWAR  
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 NUMBER OF SEQUENCES: 716  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ABBOTT LABORATORIES  
 STREET: 100 ABBOTT PARK ROAD  
 CITY: ABBOTT PARK  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500

COMPUTER READABLE FORM:

NAME: POREMBSKI, PRISCILLA E.  
 REGISTRATION NUMBER: 33,207  
 REFERENCE/DOCKET NUMBER: 5527.PC.01  
 TELEPHONE: 708-937-6365  
 TELEFAX: 708-938-2623  
 INFORMATION FOR SEQ ID NO: 387:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2972 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-488-446-387

Query Match 63.6%; Score 42; DB 4; Length 2972;  
 Best Local Similarity 62.5%; Pred. No. 7.1e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CYWGCGYW 8  
 Db 295 CQWGSAIW 302

---

RESULT 11  
 US-08-424-550B-387  
 ; Sequence 387, Application US/08424550B  
 ; Patent No. 6720166  
 GENERAL INFORMATION:  
 APPLICANT: JOHN N. SIMONS  
 APPLICANT: TAMI J. PILOT-MATIAS  
 APPLICANT: GEORGE J. DAWSON  
 APPLICANT: GEORGE G. SCHLAUDER  
 APPLICANT: SURESH M. DESAI  
 APPLICANT: THOMAS P. LEARY  
 APPLICANT: ANTHONY SCOTT MUERHOF F  
 APPLICANT: JAMES C. ERKER  
 APPLICANT: SHERI L. BULJK  
 APPLICANT: ISA K. MUSHAHWAR  
 TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 NUMBER OF SEQUENCES: 716  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
 STREET: 100 ABBOTT PARK ROAD  
 CITY: ABBOTT PARK  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/424,550B  
 FILING DATE: 07-Jun-1995  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/424,550  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: POREMBSKI, PRISCILLA E.  
 REGISTRATION NUMBER: 33,207  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 708-937-6365  
 TELEFAX: 708-938-2623  
 INFORMATION FOR SEQ ID NO: 387:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2972 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

NAME: POREMBSKI, PRISCILLA E.  
 REGISTRATION NUMBER: 33,207  
 REFERENCE/DOCKET NUMBER: 5527.PC.01  
 TELEPHONE: 708-937-6365  
 TELEFAX: 708-938-2623  
 INFORMATION FOR SEQ ID NO: 387:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2972 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-467-344A-387

Query Match 63.6%; Score 42; DB 4; Length 2972;  
 Best Local Similarity 62.5%; Pred. No. 7.1e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CYWGCGYW 8  
 Db 295 CQWGSAIW 302

---

RESULT 11  
 US-08-424-550B-387  
 ; Sequence 387, Application US/08424550B  
 ; Patent No. 6720166  
 GENERAL INFORMATION:  
 APPLICANT: JOHN N. SIMONS  
 APPLICANT: TAMI J. PILOT-MATIAS  
 APPLICANT: GEORGE J. DAWSON  
 APPLICANT: GEORGE G. SCHLAUDER  
 APPLICANT: SURESH M. DESAI  
 APPLICANT: THOMAS P. LEARY  
 APPLICANT: ANTHONY SCOTT MUERHOF F  
 APPLICANT: JAMES C. ERKER  
 APPLICANT: SHERI L. BULJK  
 APPLICANT: ISA K. MUSHAHWAR  
 TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 NUMBER OF SEQUENCES: 716  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
 STREET: 100 ABBOTT PARK ROAD  
 CITY: ABBOTT PARK  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/424,550B  
 FILING DATE: 07-Jun-1995  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/424,550  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: POREMBSKI, PRISCILLA E.  
 REGISTRATION NUMBER: 33,207  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 708-937-6365  
 TELEFAX: 708-938-2623  
 INFORMATION FOR SEQ ID NO: 387:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2972 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

Patent No. 6699703  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al.  
 TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus FILE REFERENCE: PATH00-07A  
 CURRENT APPLICATION NUMBER: US/09/583.110  
 CURRENT FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/107,433  
 PRIOR FILING DATE: 1998-06-30  
 PRIOR APPLICATION NUMBER: US 60/085,131  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: US 60/051,553  
 PRIOR FILING DATE: 1997-07-02  
 NUMBER OF SEQ ID NOS: 5322  
 SEQ ID NO: 4873  
 LENGTH: 63  
 TYPE: PRT  
 ORGANISM: Streptococcus pneumoniae  
 US-09-583-110-4873

Query Match 62.1%; Score 41; DB 4; Length 63;  
 Best Local Similarity 50.0%; Pred. No. 40;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CYWGCYW 8  
 Db 36 CFWGWSFW 43

RESULT 13  
 US-09-107-433-4664  
 Patent No. 6800744  
 GENERAL INFORMATION:  
 APPLICANT: Lynn A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 5206  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEES: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02454  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD/ROM ISO9660  
 COMPUTER: <Unknown>  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: <Unknown>  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,433  
 FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/ 085431  
 FILING DATE: MAY 12, 1998  
 APPLICATION NUMBER: 60/051553  
 FILING DATE: JULY 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Arinello, Pamela Dencke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-011  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 893-5007  
 TELEX/FAX: (781) 893-8277  
 INFORMATION FOR SEQ ID NO: 4664:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 71 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

Query Match 62.1%; Score 40; DB 1; Length 523;  
 Best Local Similarity 62.5%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CYWGCYW 8  
 Db 494 CFWGPSTW 501

RESULT 15  
US-08-483-146A-2  
Sequence 2, Application US/08483146A  
Patent No. 5696080  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
TITLE OF INVENTION: THEREFROM.  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483.146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY AGENT INFORMATION:  
NAME: Israelson, Ned A  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-483-146A-2

Query Match Score 40; DB 1; Length 523;  
Best Local Similarity 60.6%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 CYWGGYW 8  
Db 494 CIWGPSYW 501

Search completed: April 18, 2005, 20:00:40  
Job time : 23.3256 secs

This Page Blank (uspto)